GEOMETRIC NUMERICAL INTEGRATION
OF INEQUALITY CONSTRAINED, NONSMOOTH HAMILTONIAN SYSTEMS

DANNY M. KAUFMAN[†,‡] AND DINESH K. PAI[‡]

† COLUMBIA UNIVERSITY ‡ UNIVERSITY OF BRITISH COLUMBIA

Abstract. We consider the geometric numerical integration of Hamiltonian systems subject to both equality and *"hard" inequality* constraints. As in the standard geometric integration setting, we target long-term structure preservation. We additionally, however, also consider invariant preservation over persistent, simultaneous and/or frequent boundary interactions. Appropriately formulating geometric methods to include such conditions has long-remained challenging due the inherent nonsmoothness they impose. To resolve these issues we thus focus both on symplectic-momentum preserving behavior *and* the preservation of additional structures, unique to the inequality constrained setting. Leveraging discrete variational techniques, we construct a family of geometric numerical integration methods that not only obtain the usual desirable properties of momentum preservation, approximate energy conservation and equality constraint preservation, but also enforce multiple simultaneous inequality constraints, obtain smooth unilateral motion along constraint boundaries and allow for both nonsmooth and smooth boundary approach and exit trajectories. Numerical experiments are presented to illustrate the behavior of these methods on difficult test examples where both smooth and nonsmooth active constraint modes persist with high frequency.

1. Introduction. Recent attention has focused on geometric numerical integration methods that closely preserve invariants of continuous systems [Sanz-Serna and Calvo 1994; Hairer et al. 2002; Leimkuhler and Reich 2004]. In particular, symplectic-momentum preserving integration schemes have been successfully applied over a wide range of applications. In the unconstrained setting, these methods exactly preserve the symplectic form and momenta, and maintain good long term energetic behavior by approximately conserving energy (up to an additive constant) over long spans of simulation.

While extensions of symplectic methods to equality constrained systems have been extensively investigated [Hairer et al. 2002; Leimkuhler and Reich 2004], appropriately formulating symplectic methods to include inequality constraints continues to remain a challenging problem that has received limited attention in the literature. The difficulty in treating these latter cases stems from the inherent nonsmoothness imposed by such conditions. Stewart [2000], in particular, notes that, despite good conservative properties in the smooth setting, symplectic methods do not necessarily translate easily to nonsmooth problems. Direct extensions of existing symplectic methods (e.g., implicit midpoint method, Newmark methods, etc.) loose their good energetic properties and often generate surprisingly nondeterministic errors when subject to inequality constraints [Stewart 2000].

In this work we consider Hamiltonian systems subject to both equality and *inequality* constraints. As in the standard geometric integration setting, we target long-term structure preservation. In the inequality constrained setting, however, we additionally consider structure preservation over persistent, simultaneous and/or frequent boundary interactions.

With these goals in mind we focus both on symplectic-momentum preserving behavior *and* the preservation of additional structures, unique to the inequality constrained setting. In particular, we develop geometric numerical integration methods that *not only* obtain the usual desirable properties of momentum preservation, approximate energy conservation and equality constraint preservation, *but also* enforce many simultaneous inequality constraints, obtain smooth unilateral motion along constraint boundaries and allow for both nonsmooth and smooth boundary approach and exit trajectories.

1.1. Problem Statement. We consider the general case of a Hamiltonian system subject to generalized constraints that restrict system configurations, q, to an *admissible* set, A (i.e., $q \in A$). In general, this admissible set can be both nonconvex and nonsmooth.

More concretely, when the admissible set is given by a manifold, it will often be defined by a set of constraint equations of the form

$$f(q) = (f_0(q), ..., f_n(q))^T = 0. \tag{1.1}$$

For these cases the resulting holonomic Hamiltonian system reduces to a set of differential algebraic equations (DAE) for which a wide range of existing geometric approaches may be suitable [Hairer et al. 2002; Leimkuhler and Reich 2004].

In the following, we consider generalizations where the admissible set, A, can always be described by both a set of constraint equalities (as above) *and* constraint inequalities, given by functions of the configuration variables,

$$g(q) = (g_0(q), ..., g_m(q))^T \geq 0. \tag{1.2}$$

We start by considering the underlying, unconstrained system, with its *natural* Hamiltonian[1] and Lagrangian functions,

$$H(\mathsf{q},\mathsf{p}) = \frac{1}{2}\mathsf{p}^T \mathsf{M}^{-1}\mathsf{p} + V(\mathsf{q}) \quad \text{and} \quad L(\mathsf{q},\dot{\mathsf{q}}) = \frac{1}{2}\dot{\mathsf{q}}^T \mathsf{M}\dot{\mathsf{q}} - V(\mathsf{q}). \tag{1.3}$$

We can then compactly include constraint forces by adding an additional potential term that, when extremized, will penalize all nonadmissible configurations to ensure constraint compliance. In particular, to enforce so-called "hard" or "exact" constraints it is standard to consider the most extreme penalization available; we apply the extended value indicator function that *infinitely* penalizes noncompliance. For an arbitrary admissible set A, the indicator penalty function is given by

$$I_\mathsf{A}(\mathsf{x}) = \left\{ \begin{array}{l} 0, \; \mathsf{x} \in \mathsf{A} \\ \infty, \; \mathsf{x} \notin \mathsf{A}. \end{array} \right. \tag{1.4}$$

The full, nonsmooth, inequality constrained, Hamiltonian and Lagrangian formulations for the system are then obtained by augmenting the unconstrained system's natural potential function, V, with the indicator function on the admissible set, I_A [Clarke 1983].

The corresponding generalized equations of motion for the constrained system are then given by the Euler-Lagrange differential inclusion (DI),

$$\mathsf{M}\ddot{\mathsf{q}} + \nabla V(\mathsf{q}) \in -\partial I_\mathsf{A}(\mathsf{q}). \tag{1.5}$$

This inclusion, obtained from the application of ∂, the generalized gradient operator [2], to the nonsmooth penalty function [Rockafellar and Wets 1998], indicates that feasible constraint forces must lie in the inward pointing normal cone to the admissible set at q, given by $-\partial I_\mathsf{A}(\mathsf{q})$. On the interior of the feasible set the normal cone is trivially the origin. In this case the Euler-Lagrange DI simply reduces to the standard unconstrained equations of motion. Elsewhere, on the boundary, where constraint forces may be necessary, the normal cone is is given by a nontrivial, negative span of the active constraint gradients at q. See Figure 1.1.
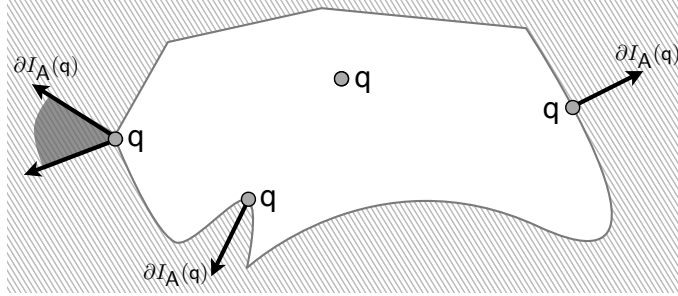

FIG. 1.1. *An example of an admissible set and a few evaluations of its associated normal cones.*

Our problem is then to derive geometric numerical schemes that integrate the generalized Euler-Lagrange system in (1.5) forward in time while preserving momenta *and* maintaining the good approximate energy preserving properties that are standard for symplectic integration schemes in the unconstrained setting [Hairer et al. 2002]. In the inequality constrained setting, however, there is additional critical behavior to consider:

1. **Smooth trajectories along the admissible set boundary.** Consider, as a simple example, a mass-particle moving with a tangential velocity across a smooth region of the boundary (Figure 1.2 (a)). If a force is pushing the particle into the boundary, the particle's trajectory should remain smooth along the boundary. No upward bounce or other type of normal motion should result. Likewise, unless dissipative forces are additionally applied, constraints should not apply "boundary capturing"-type forces that slow the particle's tangential motion. Nevertheless, existing symplectic methods generally

---

[1]Throughout, to simplify discussion and visualization, we consider the case of separable Hamiltonians with a flat metric on configuration.

[2]Note that here and in the following we reserve the ∂ notation to exclusively designate the generalized gradient operator [Rockafellar and Wets 1998].

fail in one or both of these categories. This has important implications since, in the general setting, many important physical phenomena are covered by the abstraction of smooth boundary motion. Such smooth boundary systems include oscillatory elastica resting on foundations, twisted elastic knots (such as DNA filaments), mechanical and biological grasping models, masonry structures, granular flow, as well as many other boundary relative systems that are important across structural engineering, robotics, mechanics, biology and other critical application areas.

2. **Nonsmooth exit and approach trajectories.** Many trajectories in the hard constraint context *can not be resolved by finite forces*. For instance, consider a mass-particle approaching a smooth boundary in the normal direction (Figure 1.2 (b)). To resolve the impact, the mass-particle's trajectory must be reflected about the surface normal and thus we require a nonsmooth transition. More generally, whenever a trajectory is non-tangential to the admissible set boundary, an instantaneous nonsmooth jump in state is demanded. Nevertheless many symplectic-based approaches require, by construction, that constraints always be resolved by forces. Nonsmooth motion, however, is an intrinsic aspect of many important physical models applied to better understand the complex behaviors exhibited by mechanical collisions, multi-impact systems, shock-wave propagation in granular media, molecular dynamics (using hard-sphere constraints), geophysical phenomena (e.g., earthquakes, iceberg-calving, etc.) as well as other difficult, nonlinear systems. As such, appropriately resolving nonsmooth motion is likewise critical.

3. **Smooth exit and approach trajectories.** Consider once again the simple example of a mass-particle traveling smoothly along a boundary subject to a force pushing into the boundary (Figure 1.2 (c)). If the particle encounters a sufficiently large concavity, it must be able to "break contact" and leave the surface. Likewise, if a similar mass-particle is on a ballistic trajectory, that approaches the boundary tangentially (Figure 1.2 (d)), it must not bounce back off, but instead maintain a smooth on-surface trajectory. More generally, smooth trajectories, on the admissible set boundary, should not, in any way, be constrained to remain on the boundary, while initially off-boundary trajectories, that encounter the boundary tangentially, should remain smooth.

Finally, we note that many fundamental applications of inequality constrained systems (e.g., granular systems, biomechanical locomotion, structural engineering) intrinsically *combine* both smooth and nonsmooth modes and thus require the appropriate resolution of both modes within the same framework.
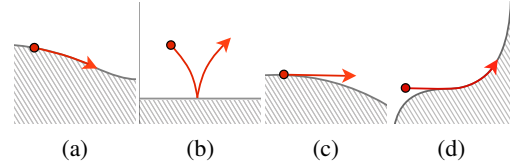

FIG. 1.2. *Examples of (a) smooth-boundary, (b) nonsmooth, (c) smooth-exiting and (d) smooth-approaching trajectories.*

1.2. Overview. In the following sections we start by discussing related approaches (§1.3) and then begin to incrementally construct our numerical integration methods, step by step. We first consider the discrete variational perspective (§2) and note that a nonsmooth, discrete Hamilton's Principle for unilateral constraints has not been previously considered. To address this gap we propose a Discrete Nonsmooth Hamilton's Principal that leads to a Discrete Euler-Lagrange Inclusion (DELI) formulation (§2.2). We then observe that standard discrete variational structure is not sufficient, on its own, to define a well-posed integrator in this setting (§2.3). The proposed DELI formulation thus provides a framework for numerical integration in addition to which further structure is required to compose a fully specified integration method. To instantiate the first such DELI-based method we next consider the structure-preserving behavior of both smooth (§3.4) and nonsmooth (§3.5) time-continuous constrained trajectories. With these observations in place, we then return to the time-discrete problem (§4), and show that introducing time-discrete analogues of smooth (§5) and nonsmooth (§6) boundary structures to DELI allows us to construct a corresponding, symplectic-momentum preserving, smooth-discrete integrator and a momentum preserving, nonsmooth-discrete integrator with observed good long-term energy behavior. Finally, in §7, we derive a discrete, generalized variational integrator (GVI) for generalized inequality-equality constrained systems that resolves both smooth and nonsmooth modes. Further numerical examples illustrating these methods are presented and discussed in §8.

1.3. Related Work. To date, approaches for the structure preserving integration of such inequality constrained, nonsmooth Hamiltonian systems can effectively be divided into two strategies:

1.3.1. Direct-Substitution Methods. Energy-momentum and symplectic-momentum methods have been extended to the inequality case by the direct substitution of a nonsmooth constraint force [Laursen and Chawla 1997; Kane et al. 1999; Stewart 2000; Laursen and Love 2002; Pandolfi et al. 2002; Deuflhard et al. 2008;
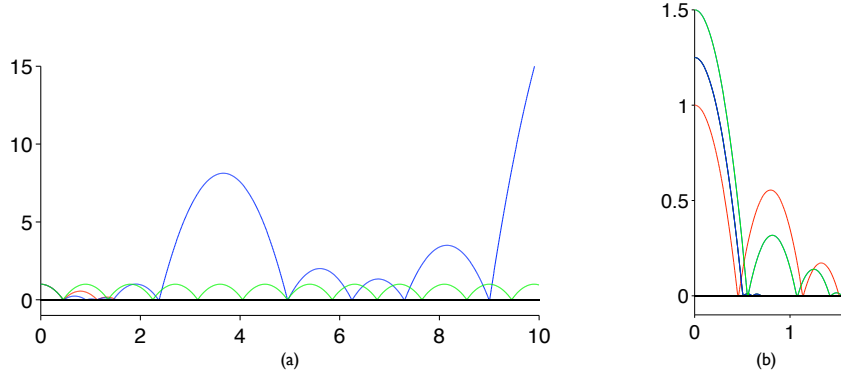
Fig. 1.3. **Point-Mass Example:** *In these figures we examine the behavior of direct substitution methods on the simple example of a one-dimensional point-mass subject to gravity and a single ground plane constraint at zero. We plot position along the y-axis and time along the x-axis. All examples shown here begin with implicit midpoint as the base, unconstrained numerical method. In* **(a)**, *starting with $q(0) = 1$, $p(0) = 0$, we plot the trajectories of the midpoint-constrained direct method in blue, the endpoint-constrained direct method in red, and the exact solution's trajectory is plotted in green. Note, as well, that all trajectories are indentical until the first impact. In* **(b)** *we plot the the trajectories of the end-point constrained direct method for three small perturbations in initial position. Note that while the endpoint-constrained direct method generates fully dissipative solutions for all cases, the degree of dissipation varies in an energetically inconsistent fashion.*

Khenous et al. 2008; Krause and Walloth 2009; Doyen et al. 2011]. Following the time continuous case discussed above, the extended value Hamiltonian for inequality-constrained systems is naturally composed by concatenating the unconstrained system's potential with the extended value indicator function. Nonsmooth extensions of standard geometric methods are then composed by direct substitution of the obtained nonsmooth constraint force, $-\partial I_A(q)$, as a standard force, handled as dictated by each numerical method. (See §B.1 for examples.)

Unfortunately, as foreshadowed above, direct substitutions generally destroy energy conservation and generate time-step dependent, nondeterministic behaviors when applied to symplectic methods. In particular, as we prove in Appendix B, *all direct-substitution, nonsmooth, one-step, symplectic methods can not, in general, preserve, either approximately, nor exactly, the energy of the Hamiltonian system.* In practice, this results in undesirable time-step-dependent and position-based restitution errors that can produce *both* energy growth *and* dissipation behaviors, while momentum similarly drifts.

Energy-momentum based direct methods and a variety of extensions and stabilizations are similarly an active and promising area of research [Doyen et al. 2011]. Although comparable behaviors to direct symplectic schemes are currently observed in practice [Doyen et al. 2011], drift and stability properties of these methods remain under investigation.

Example: A One Degree of Freedom Particle System. To illustrate some of these issues consider the following simple example of a one dimensional particle with mass, $m = 1$, configuration, $q \in \mathbb{R}$, subject to a simple unilateral constraint $g(q) = q \geq 0$, under gravity such that $V(q) = 9.8q$. (Also see Figure 1.3.) Stewart [2000] uses this simple example to illustrate some of the potential issues involved in applying direct methods to existing symplectic integration methods. Poor energy conservation behaviors are observed for this case by the author for both the (symplectic) implicit midpoint rule [Hairer et al. 2002] and the IMEX Newmark scheme from Kane et al. [1999]. In particular, Stewart [2000] notes that the energy of these systems is not conserved and that the effective coefficient of restitution varies with state for both methods.

We start by dropping the particle from a height of $q(0) = 1$. In Figure 1.3(a), we plot obtained position for a range of algorithms along the *y*-axis and time along the *x*-axis. In blue we plot the midpoint-constrained direct method; e.g., using constraint forces of the form $-\partial I_A(\frac{q^{t+1}+q^t}{2})$. This generates, within the same simulation, effective coefficients of restitution that vary widely between dissipation and unstable energy-growth. In general these effects vary the with time-step size and relative position with respect to the constraint boundary. In green we plot the exact solution. Note that both standard collision integration methods (discussed below) and our new method proposed in Section 7 generate trajectories that closely match this solution.

Finally, we also plot the endpoint-constrained direct method, e.g., using constraint forces of the form $-\partial I_A(q^{t+1})$, in red. Note that while the end-point method likewise does not preserve energy, it generates purely

dissipative energy errors. For this reason, end-point constraints and/or additional dissipative stabilizations have often been favored for direct methods in the literature [Kane et al. 1999; Pandolfi et al. 2002; Deuflhard et al. 2008; Krause and Walloth 2009; Doyen et al. 2011]. The degree of dissipation, however, varies in a non-physical manner. In particular, as with the mid-point constraint, restitution for the end-point constrained direct method is likewise dependent on time-step size and the relative position of the particle with respect to the constraint boundary. This is illustrated in Figure 1.3(b), were we plot the the trajectory of the end-point constrained direct method over small changes in initial position. Note that while these three trajectories are all clearly dissipative, the degree of dissipation continues to vary in an energetically inconsistent fashion.
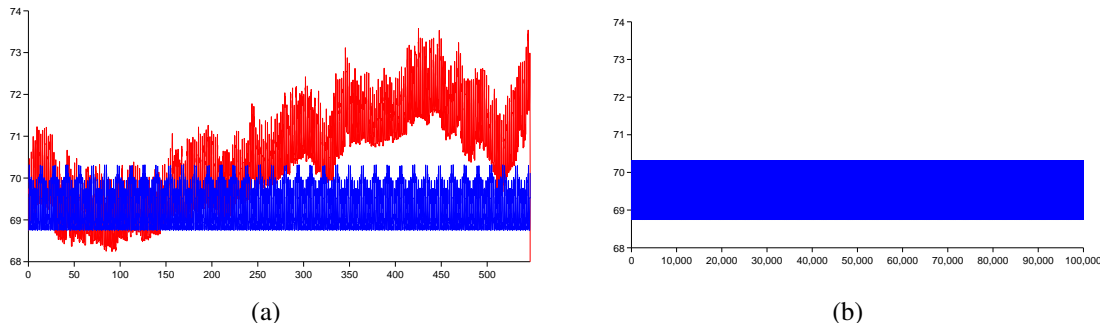

(a)   (b)

FIG. 1.4. **Pogo-Stick Example:** *In these figures we plot the energy for a simple, two-dimensional, two-point, mass-spring system vertically oriented, subject to gravity, bouncing on a ground plane constraint. Both methods investigated here begin with Störmer-Verlet as the base, unconstrained numerical method. We set $m = 1$ for both mass particles. The linear spring stiffness is set to $k = 10$, spring length is $l = 5$, and gravity is $g = 9.8213$. In (a) we plot the energy for the first 546 seconds of simulation time, stepping at $h = 10^{-1}$, for both the standard collision integration method in red and our generalized discrete variational integrator (GVI) in blue. After this step the energy in the collision integration blows up. In (b) we continue the GVI energy plot out for 100,000 seconds of simulation time.*

1.3.2. Collision Integrators. Alternately, *collision integration* methods have also been proposed [Stratt et al. 1981; McNeil and Madden 1982; Heyes 1982; Suh et al. 1990; Houndonougbo et al. 2000; Fetecau et al. 2003; Bond and Leimkuhler 2007] in which unconstrained symplectic schemes interrupt fixed-size steps each time a unique constraint is identified as becoming active (i.e., about to be violated). For each such event, an impulsive constraint force (in the form of a reflection) is computed which conserves (depending on the scheme) either a continuous [Houndonougbo et al. 2000], discrete [Fetecau et al. 2003] or numerical Hamiltonian [Bond and Leimkuhler 2007]. These methods thus all effectively apply adaptive time-step subdividivision to resolve constraints. Given these variable-step modifications, a trajectory-shadowing, numerical Hamiltonian system can no longer be defined and thus energy conservation can not be expected in the long-term [Hairer et al. 2002].

A novel stabilization scheme, based on the application of two complementary strategies, was recently proposed to improve the energetic behavior of collision integrators [Bond and Leimkuhler 2007]. Stabilization is primarily obtained by applying reflections that preserve a second order approximation of the selected integration scheme's (generally Störmer-Verlet's) numerical Hamiltonian. A secondary improvement is then also obtained by stepping to and away from constraint events using consistent, higher-order integration schemes. This latter modification, however, is motivated by applications in which constraint events are infrequent. The cost of higher-order integration is thus ameliorated by the assumption that the majority of time-steps are taken constraint free using the base, lower order method. In our development we now consider highly constrained systems where we expect most, if not all, steps to encounter inequality constraints. In this context, any potential stabilization advantage is then given almost exclusively by the higher-order energy reflection. We will examine this further in §6.1.1.

While these developments can improve the performance of collision integrators, they still demand that only a single constraint be active at any time. Moreover, all collision integrators require expensive root-finding routines to determine each such constraint event. Many applications (e.g., surgical simulation, structural engineering, robotic grasping, granular flow, etc.), however, demand the simultaneous resolution of potentially large numbers of simultaneously active constraints. Individual treatment of constraints in these contexts can

lead to ordering bias, unacceptable stability issues, and generally have an especially large detrimental impact in implicit methods. Even when such ordering issues are acceptable, (e.g., small time-step, explicit methods) the individual time-ordered evaluation of all active constraint events will be, as Cirak and West [2005] note, computationally intractable for many simulation problems (consider, for example, Zeno's paradox).

As a generalization it is reasonable, however, to expect that extensions of collision integration methods should be possible, even if not practical, which simultaneously resolve many active constraints. Indeed, one such possible generalization is suggested in Fetecau et al. [2003]. Even in this paradigm, however, the inter-leaving of fixed-step symplectic-momentum methods with a variable-step, energy method still generates drift.

To understand the practical implications of these issues in applying collision integration methods, consider the energy behaviors observed in Figure 1.4(a), above, as well as Figures 6.1(a) and 8.1(b) and the related discussion in §6.1 and §8.0.1.

2. Variational Integration and DELI Integrators. Despite the above negative results, nonsmoothly constrained Hamiltonian systems *are* invariant preserving in the usual senses [Kane et al. 1999]. Thus it is still desirable to find geometric integration methods that *do not* introduce numerical artifacts into simulations (dissipative or otherwise). One fruitful direction of research, that we do not pursue here, is the ongoing investigation in extending the discrete *energy-momentum* conserving framework to the nonsmooth setting. Instead we pursue the complementary research program of extending *symplectic-momentum* preserving schemes to resolve inequality constrained dynamics. In particular, observing the natural variational structure implied by inequality constraints [Rockafellar and Wets 1998], we revisit the application of *Discrete Variational Integration* [Marsden and West 2001] to the inequality constrained setting.

To date, two approaches towards extending Variational Integrators to the nonsmooth, inequality constrained setting have been considered:

In the first approach, *Nonsmooth Variational Integrators* are generated [Kane et al. 1999; Pandolfi et al. 2002] by extremizing the standard discrete-action obtained directly from the quadrature of the underlying, smooth, unconstrained system. This obtains a standard Variational Integrator into which a nonsmooth constraint force is then directly substituted. We observe that this approach then obtains nothing other than a direct-substitution symplectic method and is thus covered by Theorem B.1 and subject to the same drift and stability issues as discussed above in §1.3.1. Moreover, we observe that this approach violates the fundamental "discretize-first, extremize-second" mantra of discrete variational methodology.

In the second approach, Fetecau et al. [2003] observe that nonsmoothness in the variational formulation can be avoided altogether by adding an additional degree-of-freedom that parameterizes time of collision. Composing and extremizing a corresponding discrete variational principle then generates *Variational Collision Integrators* which can be shown to be symplectic [Fetecau et al. 2003]. As observed by Bond and Leimkuhler [2007], however, resulting schemes exactly obtain a standard collision integrator, as discussed above in §1.3.2, composed so that each collision preserves a discrete (rather than the continuous or numerical) energy at each reflection.

With these observations in place we then note that, to the best of our knowledge, a *nonsmooth, discrete Variational Principle for inequality constrained systems remains unexplored*. We thus begin our development of nonsmooth geometric methods by developing these initial steps in the fully "discretize-first" manner. In particular, we apply Variational Integration (VI) methods [Marsden and West 2001; Hairer et al. 2002] as a prescriptive approach for generating a family of geometric methods for inequality constrained dynamics.

2.1. VI Methods. Recall that VI methods, mimicking the continuous derivation of the Lagrangian equations of motion, extremize a discrete action on the time interval $t \in [0, T]$, over all possible discrete paths of the form $\{q^0, ..., q^k, ..., q^N\}$, where $N = T/h$. We begin with the discrete Lagrangian

$$L_d(q^k, q^{k+1}) \simeq \int_k^{k+1} L(q, \dot{q}) dt, \tag{2.1}$$

that gives a quadrature of the natural Lagrangian over a finite interval.

The sum of the discrete Lagrangians, over all intervals, then gives the corresponding standard discrete action,

$$\sum_{k=0}^{N-1} L_d(q^k, q^{k+1}). \tag{2.2}$$

Extremizing this associated discrete action then generates a unique, symplectic-momentum preserving integration scheme [Hairer et al. 2002], with accuracy specified by the order of the quadrature applied [West 2004].

2.2. DELI. As a point of simple yet, to the best of our knowledge, novel departure we now add the nonsmooth penalty term to the discrete action and formulate a discrete, nonsmooth, constrained Hamilton's Principle that enforces hard inequality constraints on all endpoints of the discrete trajectory,

$$\delta_k \sum_{k=0}^{N-1} \left(L_d(\mathsf{q}^k, \mathsf{q}^{k+1}) - I_{\mathsf{A}}(\mathsf{q}^{k+1}) \right) \ni 0. \tag{2.3}$$

Stationarity then gives us our constrained, Discrete Euler-Lagrange Inclusion (DELI),

$$D_2 L_d(\mathsf{q}^{t-1}, \mathsf{q}^t) + D_1 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}) - \partial I_{\mathsf{A}}(\mathsf{q}^t) \ni 0, \tag{2.4}$$

$$\mathsf{q}^{t+1} \in \mathsf{A}. \tag{2.5}$$

Given the last two configurations, q^t and q^{t-1}, the DELI integrator advances state to a new configuration, q^{t+1}. Here, in this recursive form, we observe that discrete constraint forces, applied to enforce constraints at the end of time-steps, are generated by normal cones constructed at the beginning of time-steps.

We can further clarify this relationship by applying a momentum matching argument [West 2004] (or, if preferred, a discrete Legendre-Fenchel transform) to obtain the corresponding discrete momentum variables, p^t and p^{t+1}, as well as the discrete phase-space map composed of a position-map inclusion,

$$D_1 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}) + \mathsf{p}^t - \partial I_{\mathsf{A}}(\mathsf{q}^t) \ni 0, \tag{2.6}$$

$$\mathsf{q}^{t+1} \in \mathsf{A}, \tag{2.7}$$

and a corresponding momentum-map update equation,

$$\mathsf{p}^{t+1} = D_2 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}). \tag{2.8}$$

A one-step formulation of the DELI then follows. Given the discrete phase-space pair, $(\mathsf{p}^t, \mathsf{q}^t)$, at time t, we now apparently can solve the forward position map system, (2.6) and (2.7), for the new constrained configuration q^{t+1}. Then, with the new configuration in hand, the discrete momentum update follows directly from (2.8).

2.3. The Underdetermined Forward Map. We are now left with an interesting question of causality with respect to constraints. In our above DELI we require that corrective forces, that appear to be generated by the configuration at the beginning of the time-step (i.e, at q^t), enforce constraints on the configuration at the *end* of the time step.

In the direct methods we discussed above in Section 1.3.1, constraint forces are instead generated by configurations defined either somewhere in the middle of the time-step, i.e. $\mathsf{q}^t + \alpha \mathsf{q}^{t+1}, \alpha \in (0,1)$, or else at the end, i.e., q^{t+1}. When solving inclusions with these terms, the variational structure of the generalized gradient then *entirely* specifies the constraint force. This, in turn, leads to the destruction of the invariant preserving properties of the underlying unconstrained method. See Appendix B, for further discussion.

To retain the geometric properties of unconstrained methods we thus apparently require some additional degree(s) of freedom in choosing our constraint forces. Due to the odd time-causality noted above, the variational structure of our DELI formulation no longer fully specifies constraint forces and thus provides such latitude for their selection. To see this, we first observe that in the DELI system merely requires constraint forces to lie in the span of the normal cone generated by the configuration at time t, i.e, $-\partial I_{\mathsf{A}}(\mathsf{q}^t)$. Since this cone is independent of final configuration, (2.6) only specifies the span of possible constraint force directions but does not, unlike in direct-substitution methods, pin down a corresponding force magnitude.

Thus, in the inequality constrained setting, we observe that standard discrete variational structure is not sufficient, on its own, to define a well-posed integrator. We find that constraint forces must be applied along the directions positively spanned by the normal cone and, likewise, all inequalities must be satisfied, but fundamentally, nothing further is specified. In particular, an underdetermined system is composed with any number of solutions, most of which will not generate symplectic-momentum preserving maps. *Our proposed DELI formulation thus provides a framework for numerical integration in addition to which further structure is required to compose a fully specified integration method.*

In the following sections we will propose the first instantiation of a DELI-based method. We specifically note that the dichotomy between nonsmooth and smooth trajectories is not well-resolved by the discrete variational framework. We will formulate our DELI-based integrator by first considering both the structure of smooth trajectories and the corresponding nonsmooth case. We will then show how imposing discrete analogues of these nonstandard structures upon the DELI formulation composes a fully defined DELI-based integration method and consider the behavior obtained by their application.

3. Time-Continuous Setting. In order to consider how to design such methods, we first focus on the ways in which the flow of the constrained Hamiltonian system is transported in the time-continuous setting.

3.1. Tangency and Normality. As a first step we introduce a few useful notational conventions and then briefly discuss variational structures that allow us to generalize notions of tangency and normality for cases involving multiple active constraints.

3.1.1. Notation. For convenience, throughout the remaining sections we adopt a constraint subset notation. Presuming, as above, that the a full constraint set is indexed by $i \in \{0, ..., m\}$, we then define, for each index subset, $\mathbb{K} \subset \{0, ..., m\}$, the corresponding constraint subset,

$$\mathsf{g}_{\mathbb{K}}(\mathsf{q}) \stackrel{\text{def}}{=} \left(\mathsf{g}_{k_0}(\mathsf{q}), ..., \mathsf{g}_{k_l}(\mathsf{q})\right)^T, \quad \{k_0, .., k_l\} \equiv \mathbb{K}, \tag{3.1}$$

constraint subset gradient.

$$\mathsf{G}_{\mathbb{K}}(\mathsf{q}) \stackrel{\text{def}}{=} \left(\nabla \mathsf{g}_{k_0}(\mathsf{q})^T, ..., \nabla \mathsf{g}_{k_l}(\mathsf{q})^T\right)^T, \quad \{k_0, .., k_l\} \equiv \mathbb{K}, \tag{3.2}$$

and the active constraint subset gradient,

$$\mathsf{N}_{\mathbb{K}}(\mathsf{q}) \stackrel{\text{def}}{=} \mathsf{G}_{\mathbb{A}\mathbb{K}}(\mathsf{q}), \quad \mathbb{A}\mathbb{K} = \left\{i : \mathsf{g}_i(\mathsf{q}) = 0, \ i \in \mathbb{K}\right\}. \tag{3.3}$$

Consistent with the notation introduced in Section 1.1 for the vector-valued constraint function, $\mathsf{g}(\mathsf{q})$, an absence of subscripting indicates that the entire constraint set is considered. We then have $\mathsf{G}(\mathsf{q}) = \mathsf{G}_{\{0,...,m\}}(\mathsf{q})$ and $\mathsf{N}(\mathsf{q}) = \mathsf{N}_{\{0,...,m\}}(\mathsf{q})$.

3.1.2. Normal and Tangent Cones. The nonsmooth generalization of a normal, the *inward pointing normal cone* [Rockafellar 1970; Clarke 1983; Rockafellar and Wets 1998] to A at q, can, in our context, be defined directly with respect to the gradients of the constraint functions, g. It is given, with respect to the full active constraint gradient, $\mathsf{N}(\mathsf{q})$, as $\{\mathsf{N}(\mathsf{q}) \, \mathsf{x} : \mathsf{x} \geq 0\}$. Alternately, we note that for q in A, the generalized gradient of the indicator function generates the *outward* pointing normal cone to the feasible set. Thus, for all configurations we have

$$\partial I_{\mathsf{A}}(\mathsf{q}) = \{-\mathsf{N}(\mathsf{q}) \, \mathsf{x} : \mathsf{x} \geq 0\}. \tag{3.4}$$

We note that, in particular, the normal cone reduces, on the interior of A, to the singleton, $\{0\}$. Correspondingly, recalling the Euler-Lagrange DI, given in (1.5), trajectories on the interior of the admissible set remain governed by the standard Euler-Lagrange equations of motion.

Polar to the normal cone, *the tangent cone* to A at q, can also be defined with respect to the active constraint gradients, N, as

$$\mathsf{T}(\mathsf{q}) \stackrel{\text{def}}{=} \{\mathsf{y} : \mathsf{N}(\mathsf{q})^T \mathsf{y} \geq 0\}. \tag{3.5}$$

When the admissible set is locally regular, the tangent cone, $\mathsf{T}(\mathsf{q})$, encodes the set of feasible directions at q, along which infinitesimal motion is locally permissible while, more generally, as we will discuss below, tangent cones generate useful generalizations of tangent spaces to nonsmooth regions of the boundary.

3.2. Constrained Trajectories. In the inequality-constrained setting there is a critical dichotomy between nonsmooth and smooth motion. Consider the simple example of a mass-particle constrained to lie in a nonsmooth valley illustrated in Figure 3.1 (a). If the particle lies on the valley floor, then it could have arrived via some motion that is locally tangent to the valley's boundaries, such as moving along the valley floor, e.g., the green arrow in Figure 3.1 (a). In this case smooth trajectories, such as continuing along the valley floor, exist
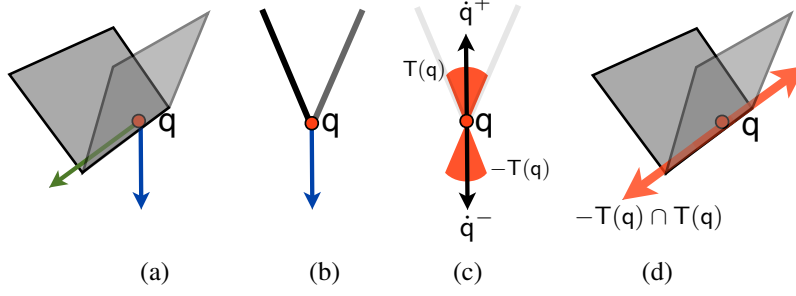
FIG. 3.1. *A simple motivating example of smooth and nonsmooth trajectories for constrained systems. See §3.2 and §3.4 for discussion.*

that can continue to satisfy admissibility. Alternately, the particle could have arrived at the valley floor with a downward pointing velocity, e.g., the blue arrows in Figure 3.1 (a) and in the cut-away view in Figure 3.1 (b). In this case there are no possible smooth trajectories that can continue to satisfy admissibility. Some sort of nonsmooth jump at the boundary is then required.

To allow for the possibility of nonsmooth jumps we consider the left and right limits of velocity,

$$\dot{\mathsf{q}}^-(t) = \lim_{\tau\uparrow t}\dot{\mathsf{q}}(\tau) \quad \text{and} \quad \dot{\mathsf{q}}^+(t) = \lim_{\tau\downarrow t}\dot{\mathsf{q}}(\tau), \tag{3.6}$$

where we assume that $\dot{\mathsf{q}}$ has bounded variation on finite time intervals; limits on momentum follow similarly. We then consider trajectory intervals and events of interest with respect to the constraint set:

DEFINITION 1. *(Smooth Trajectory Intervals and Nonsmooth Trajectory Events). We identify* **smooth trajectory intervals** *as time intervals* $[a,b] \subseteq \mathbb{R}$ *where* $\dot{\mathsf{q}}(t) = \dot{\mathsf{q}}^-(t) = \dot{\mathsf{q}}^+(t)$ *for all* $t \in [a,b]$ *and* **nonsmooth trajectory events** *as times* $t \in \mathbb{R}$ *where* $\dot{\mathsf{q}}^-(t) \neq \dot{\mathsf{q}}^+(t)$.

3.3. Kinematic Conditions. Now consider the cut-away view of our point-mass example again, shown in Figure 3.1 (c) with both the tangent and negated tangent cones depicted in red. If the mass-particle lies on the boundary at time *t*, then it must have approached the boundary (in the left time limit) with a velocity in the negated tangent cone, since no other possible trajectory exists that could have reached the boundary. Likewise, in order to remain an admissible trajectory, the particle must leave the boundary (in the right time limit) with a velocity that is non-negative with respect to all active constraint gradients; thus the right-limit velocity should lie in the positive tangent cone.

This intuition is made rigorous by a Lemma due to Moreau [1988, Proposition 2.2], which says that a trajectory lies in the admissible region *for all time* if and only if left velocities always lie in the negated tangent cone and right velocities lie in the positive tangent cone:

LEMMA 3.1.

$$\mathsf{q}(t) \in \mathsf{A}, \ \textit{for all } t \iff \dot{\mathsf{q}}^-(t) \in -\mathsf{T}(\mathsf{q}) \ \textit{and } \dot{\mathsf{q}}^+(t) \in \mathsf{T}(\mathsf{q}), \ \textit{for all } t. \tag{3.7}$$

For the nonsmooth case, as in the simple example discussed above, we expect that the left limit will not equal the right and thus we encounter a nonsmooth event requiring an instantaneous jump. We will discuss the time-continuous, nonsmooth case more thoroughly below in Section 3.5. First, however, we consider the time-continuous, smooth case in greater detail.

3.4. Smooth Trajectories. We note that smooth motion, with respect to inequality constraints, corresponds at all times to either (unconstrained) motion on the interior of the admissible set, or to any combination of the following boundary cases: (1) encountering constraint boundaries tangentially, (2) leaving constraint boundaries, or (3) tangential motion along constraint boundaries.

More concretely, smooth trajectory intervals are characterized by a *tangent subspace condition*:

LEMMA 3.2. *If* $[a,b]$ *is a smooth trajectory interval then*

$$\dot{\mathsf{q}}^-(t) = \dot{\mathsf{q}}^+(t) \in -\mathsf{T}(\mathsf{q}(t)) \cap \mathsf{T}(\mathsf{q}(t)), \ \textit{for all } t \in [a,b]. \tag{3.8}$$

9

Proof. Follows directly from Lemma 3.1 and Definition 1. □

The intersection of the tangent cone and the negated tangent cone thus defines the *tangent subspace* at nonsmooth portions of the admissible set boundary (see our example Figure 3.1 (d)), along which smooth motion is possible and trivially generates the tangent space to smooth regions of the boundary. Conversely, we note that if the left velocity lies in the tangent subspace then, locally, an admissible smooth trajectory exists.

In particular, while the structure of the normal cone implies that the standard (pointwise) Signorini-Fichera [Kikuchi and Oden 1988] complementarity[3] condition, $g(q) \perp \lambda$, holds everywhere, over smooth trajectory intervals Lemma 3.2 and the Euler-Lagrange DI in (1.5) imply that a stronger *constraint-force continuity condition* also holds per constraint:

THEOREM 3.3. *Let* $g_i\big(q(a)\big) = 0$ *and* $\nabla g_i\big(q(a)\big)^T \dot{q}^-(a) = 0$; *if* $\lambda_i(t) > 0$ *holds for all* $t \in [a,b]$ *then*
(a) $g_i\big(q(t)\big) = 0$, *and*
(b) $\nabla g_i\big(q(t)\big)^T \dot{q}^+(t) = 0$,
for all $t \in [a,b]$.

Proof. We can equivalently define the inward-pointing normal cone with respect to *all* constraint gradients as $G(q)\lambda = -\partial I_A(q)$ with $g_i(q) > 0 \implies \lambda_i = 0$ and $\lambda_i > 0 \implies g_i(q) = 0$. In turn, this implies that if $\lambda_i(t) > 0$ on an interval, i.e., $t \in [a,b]$, we must then correspondingly have $g_i(q(t)) = 0$ for all $t \in [a,b]$. The chain rule then gives $\nabla g_i\big(q(t)\big)^T \dot{q}^+(q(t)) = 0$ on the same interval. □

3.5. Nonsmooth Motion. With these observations on smooth motion in place, we turn to consider nonsmooth motion. As discussed above, when new constraints are encountered non-tangentially, i.e., from a normal direction, a nonsmooth trajectory event occurs (see Figure 3.1 (c)). In the general setting this corresponds to the case where the left limit velocity lies strictly in the interior of the negated tangent cone and requires the satisfaction of *jump conditions*:

THEOREM 3.4. *If* $\dot{q}^-(t) \notin T(q)$ *then*
(a) $\dot{q}^+(t) \neq \dot{q}^-(t)$,
(b) $p^+(t) = p^-(t) + dp$,
(c) $dp \in -\partial I_A(q)$, *and*
(d) $\dot{q}^+(t) \in T(q)$.

Proof. By Lemma 3.1 since $\dot{q}^-(t) \notin T(q)$, $\dot{q}^-(t)$ must lie strictly in the interior of $-T(q)$ and, again by Lemma 3.1, right feasibility satisfaction of $\dot{q}^+(t) \in T(q)$ then requires a nonsmooth trajectory event at time t where $\dot{q}^+(t) \neq \dot{q}^-(t)$. To satisfy right feasibility, a jump, dp, in momentum is required to obtain $p^+(t) = p^-(t) + dp \in MT(q)$. Finally, at t, the Euler-Lagrange DI reduces [Brogliato 1999] to the measure-valued, jump inclusion, $dp \in -\partial I_A(q)$, which fully determines the feasible set of possible jumps. □

At nonsmooth trajectory events, however, although the measure-valued Euler-Lagrange system, $dp \in -\partial I_A(q)$, continues to enforce momentum conservation, it no longer specifies energy behavior. Conservation of energy is then *explicitly* invoked by the side condition

$$H(q, p^+) = H(q, p^-). \tag{3.9}$$

The combined criteria of the *jump conditions* and *energy conservation* are satisfied by all conservative, nonsmooth jumps at the boundary of the admissible set. In particular, these jump conditions uniquely define an energy and momentum preserving reflection whenever a single constraint is considered. In the general, multi-constraint setting, however, although the jump conditions effectively delimit the space of solutions, an invariant-preserving, constraint-enforcing, impulse is no longer uniquely given [Glocker 2004]. In these cases, jumps can effectively be resolved by any of an infinite set of feasible corrective impulses.

We note, however, that a geometric treatment of inequality constrained systems requires a well-posed resolution of jumps with respect to multiple simultaneously active constraints. Methods that satisfy well-posedness by providing uniquely defined, physically motivated, multi-impact solutions for the multi-constraint setting are thus desirable [Chatterjee and Ruina 1998]. In particular, in the following examples, we employ the generalized reflection operator of Kaufman et al. [2010]. Alternately, the multi-impact solutions of Moreau [1988] or Chatterjee and Ruina [1998] can also be applied.

---

[3]Note that here and in the following we are using the standard complementarity notation, $x \perp y$, to indicate $x_i y_i = 0$ holds componentwise for the matching entries of $x, y \in \mathbb{R}^{m+1}$.

4. Discrete Setting. Given the above analysis, we observe that nonsmoothly constrained Hamiltonian systems correspond to smooth Hamiltonian systems almost everywhere. At instants of transition, where new constraints are encountered or left behind, *either* a smooth trajectory is maintained across the transition *or* a jump provides a nonsmooth mapping between two consistent smooth trajectories. For smooth transitions, motion is given by the Euler-Lagrange DI while, for nonsmooth transitions, jumps instantaneously reflect to a new, energetically consistent, symmetry preserving, smooth Hamiltonian flow. As discussed above in §3.4 and §3.5, the conditions specifying the mode of each such transition are entirely given by the relationship between the left limit velocity and the tangent cone.

To carry these observations into the discrete setting, we will thus first consider compatible, discrete-predicates for distinguishing between smooth and nonsmooth transition modes at each time-step. With these predicates in place, we will then generate compatible discrete analogues of the time-continuous, smooth and nonsmooth maps.

As a prescriptive guide for constructing these maps, we next observe that, *if* we require our discrete maps to reduce to symplectic steps in the unconstrained case, any obtained *unconstrained* step ending *on* the constraint boundary is always trivially shadowed by an unconstrained numerical Hamiltonian system [Hairer et al. 2002]. If the successive *constrained* time-step is *discrete-smooth*, then we should require the corresponding discrete-smooth map to be *symplectic*, so that it is likewise consistently shadowed. Similarly, if a successive constrained time-step is *discrete-nonsmooth*, then the corresponding discrete-nonsmooth jump-map should be constructed so that it transitions to a new, uniquely defined, invariant preserving numerical Hamiltonian system that is similarly consistent. Following this guide, we conjecture that a piecewise-smooth, numerical Hamiltonian system, for the full inequality constrained system, can then be constructed by gluing together the corresponding smooth numerical Hamiltonians over successive smooth and nonsmooth transitions.

Below we will employ the preceding time-continuous analysis to show one way in which such discrete maps can be generated. The resulting integrator will, by construction, attempt to preserve correspondence to such a composite, shadowing Hamiltonian system. We note that one fundamental requirement of maintaining such a correspondence is that all new constraint boundaries be considered, in the discrete setting, only at the end (and/or beginning) of each time-step. Indeed, it is specifically the explicit violation of this particular criterion that causes collision integrators to perform so poorly in the nonsmooth case[4]. With these observations in place we now begin the development of our proposed, DELI-based, nonsmooth geometric methods.

5. Discrete Smooth Setting. In the discrete setting, smooth-trajectory intervals are determined by considering whether the discrete-analogue of the *tangent subspace condition* in (3.8) is satisfied at interval endpoints along the discrete path such that

$$\mathsf{M}^{-1}\mathsf{p}^t \in -\mathsf{T}(\mathsf{q}^t) \cap \mathsf{T}(\mathsf{q}^t). \tag{5.1}$$

We then identify each DELI forward-map with a discrete-smooth interval on the discrete path. Making all constraint functions and their gradients explicit in (5.1) we find that the discrete left-limit

$$g_i(\mathsf{q}^t) = 0 \iff \nabla g_i(\mathsf{q}^t)^T \mathsf{M}^{-1}\mathsf{p}^t = 0, \tag{5.2}$$

and discrete right-limit

$$g_i(\mathsf{q}^{t+1}) = 0 \iff \nabla g_i(\mathsf{q}^{t+1})^T \mathsf{M}^{-1}\mathsf{p}^{t+1} = 0, \tag{5.3}$$

subspace tangency conditions simply require momentum to lie in the cotangent spaces of all active constraints at the beginning and end of all smooth intervals.

The discrete-analogue of the *constraint-force continuity condition* in Theorem 3.3 is then similarly given by

$$\lambda_i > 0 \implies g_i(\mathsf{q}^{t+1}) = 0. \tag{5.4}$$

5.1. Discrete-Smooth Integrator. Reconsidering our derived DELI in (2.6) we first make constraint functions and their corresponding gradients explicit to obtain the equivalent forward map system,

$$
\begin{aligned}
D_1 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}) + \mathsf{p}^t + \mathsf{N}(\mathsf{q}^t)\lambda &= 0, \quad \lambda \geq 0, \\
\mathsf{g}(\mathsf{q}^{t+1}) &\geq 0, \\
\mathsf{p}^{t+1} &= D_2 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}).
\end{aligned}
\tag{5.5}
$$

---

[4]Of course, by construction, collision integrators are explicitly designed only for the nonsmooth case.
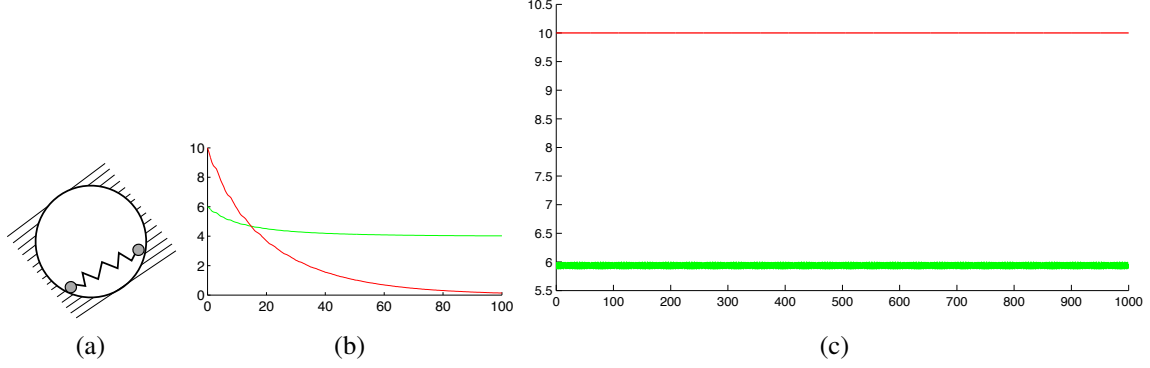
FIG. 5.1. *As illustrated in (a), we consider a smooth trajectory of a constrained system composed of a simple, four-dimensional, two-point, mass-spring system, subject to a circle inequality constraint and an outward-pulling, radial potential. We initialize both particles to a rotational trajectory, initially tangent to the boundary and note that all constraints and potentials, in this example, are rotationally invariant and thus angular momentum should be preserved. Both methods investigated here begin with implicit-midpoint as the base, unconstrained numerical method and are stepped at $h = 5 \times 10^{-1}$. In (b) we plot the energy (in green) and the angular momentum (in red) of the direct, end-point constrained method for the first 100 units of simulation time. We note the decay in both angular momentum and energy. In (c) we similarly plot the (approximate) preservation of energy (in green) and (exact) conservation of angular momentum (in red) obtained by our Discrete Smooth Integrator out to 1000 units of simulation time. See §5.2 below for further discussion.*

We then augment DELI with discrete left-limit tangency as a precondition, discrete right-limit tangency as a postcondition and discrete constraint-force continuity over each smooth interval to obtain a DELI-based, *Discrete-Smooth Integrator*:

$$D_1 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}) + \mathsf{p}^t + \mathsf{N}(\mathsf{q}^t)\lambda = 0, \tag{5.6}$$

$$0 \leq \lambda \perp \mathsf{g}(\mathsf{q}^{t+1}) \geq 0, \tag{5.7}$$

$$\mathsf{p}^{t+1} = D_2 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}) + \mathsf{N}(\mathsf{q}^{t+1})\,\mu, \tag{5.8}$$

$$\mathsf{N}(\mathsf{q}^{t+1})^T \mathsf{M}^{-1} \mathsf{p}^{t+1} = 0. \tag{5.9}$$

An update step for the Discrete-Smooth Integrator is then generated by first solving the position update equations, (5.6) and (5.7), for q^{t+1} and λ. A momentum update then follows from solving equations (5.8) and (5.9), for p^{t+1} and μ.

We then observe that

THEOREM 5.1. *The Discrete Smooth Integrator given in (5.6) – (5.9) is symplectic-momentum conserving.*
Proof. See §C.

By construction, the method likewise preserves the subspace tangency of discrete smooth trajectories when moving along, leaving and approaching constraint boundaries. Likewise, the integrator trivially reduces to the unconstrained symplectic method, given by the discrete Lagrangian, on the interior of the admissible set.

5.2. Spring and Sphere Example. As an initial consideration of the Discrete Smooth Integrator and in comparison to the direct-substitution schemes discussed in §1.3.1, we examine the behavior of a simple, smooth trajectory system. We consider the simple example of a four-degree of freedom, mass-spring system. Configuration is given by the positions of two mass particles, $\mathsf{q} = (\mathsf{q}_1, \mathsf{q}_2) \in \mathbb{R}^4$, connected by a spring. The system is subject to a standard spring potential between particles and an outward pulling, radial potential, i.e., $V(\mathsf{q}) = \sum_{i=1}^{2} 25/(\mathsf{q}_i^T \mathsf{q}_i) + \frac{1}{2}(\| \mathsf{q}_1 - \mathsf{q}_2 \| - l)^2$, with $l = 2\sqrt{2}$. In addition, we apply inequality constraints that require the particles to lie in a sphere, $\mathsf{g}_i(\mathsf{q}_i) = \| \mathsf{q}_i \| - r \geq 0$, of radius $r = 5$. See Figure 5.1 (a). Note that all constraints and potentials, in this example, are rotationally invariant and thus angular momentum should be preserved by the system.

We initialize the example system to a smooth, rotational trajectory: both particles start, at time $t = 0$, on the sphere boundary, $\mathsf{q}_1(0) = (4, -3)$, $\mathsf{q}_2(0) = (3, -4)$, with corresponding unit length momenta that point counter-clockwise and tangent to the boundary. This initiates an on-boundary, counter-clockwise rotation of the particles, as well as an oscillatory motion between them. Both methods investigated here begin with implicit-midpoint as the base, unconstrained numerical method and are stepped at $h = 5 \times 10^{-1}$.

In Figure 5.1 (b) we plot the energy (in green) and the angular momentum (in red) of the direct, end-point constrained method for the first 100 units of simulation time. We note that the decay in both angular momentum

12

and energy for the direct method corresponds to the destruction of both the rotational and oscillatory modes of the system.

In Figure 5.1 (c) we similarly plot the (approximate) preservation of energy, in green, and the (exact) conservation of angular momentum, in red, obtained by our Discrete Smooth Integrator out to 1000 units of simulation time. Throughout this simulation both particles maintain the smooth, on boundary, rotational and oscillatory modes of the system.

Finally we note that collision integrators are not suitable for application in such examples since they are, as discussed above, formulated to resolve only nonsmooth boundary interactions. Nevertheless if we do apply a collision integration scheme in this smooth setting (effectively applying a micro-collision model [Brogliato 1999]), we observe that while energy behavior is not compelling (see §6.1.1 below for a similar example) the trajectory is also qualitatively wrong; the nonsmooth reflections applied by the collision integrator introduce an artificial, nonsmooth, normal, bouncing mode (off the sphere boundary) that increases over time and eventually drives the system to instability.

6. Discrete Nonmooth Setting. If the discrete left smoothness predicate is *not* satisfied at time t, the left limit velocity must lie strictly in the negated tangent cone such that $\mathsf{M}^{-1}\mathsf{p}^t \notin \mathsf{T}(\mathsf{q}^t)$. A nonsmooth jump is then required to satisfy admissibility. Following our discussion in §3.5, we apply discrete, generalized, energy preserving reflections to resolve these cases. Later, in §6.3, we will focus on *how* we derive such discrete generalized reflection operators. First, however, we will consider the critical issue of *when* such reflections should be applied.

6.1. Extended Reflections. We will, in this section, temporarily ignore smooth motion and presume that all encounters with the admissible set boundary are nonsmooth and are always resolved by the application of some energy preserving, generalized reflection (to be defined below in §6.3). Presuming that we apply an integrator that reduces to a symplectic method in the unconstrained case, standard backwards error analysis [Hairer et al. 2002] guarantees that, away from boundaries, there always exists a trajectory shadowing numerical Hamiltonian. In an ideal setting each such unconstrained, symplectic time-step would end either in the interior or *exactly* on the boundary of the admissible set.

Correspondence to the piecewise-smooth, numerical Hamiltonian, discussed above in §4, then follows. Of course, such nice behavior can not be expected in our generally unsynchronized universe. Instead, we are faced with the likelihood that most, if not all, new constraint boundaries will be encountered midstep.

The standard approach of existing collision integration schemes (see §1.3.2) is to take unconstrained symplectic steps until the admissible boundary is crossed. When a boundary is crossed, collision integrators return to the beginning of the time-step and then repeat the forward step using *just* the time-step fraction that lands configuration exactly on the active boundary. A energy-preserving reflection is then performed using current state and then, finally, the integrator completes the remaining portion of the full time-step.

A simple alternate approach, that, to the best of our knowledge, we introduce for the first time here, is to *always take full time-steps*, using augmented or *extended reflections* to incorporate otherwise missed constraint events. More concretely, we apply smooth symplectic steps whenever the admissible boundary is not encountered from a normal approach. Anytime current configuration, q^t, indicates that a constraint is active and/or a predictor configuration, q^p, indicates that a constraint violation is imminent or alternately, depending on the choice of predictor, has already occurred, we perform a generalized reflection at the beginning (equivalently end) of the time step, with respect to both active constraints and *predicted* active constraints. Once the reflection is applied we then take the full smooth, symplectic step. This ensures correspondence between numerical Hamiltonians (as discussed above in §4), at the cost of inducing errors in the representation of the local constraint boundary geometry.

6.1.1. Pogo-Stick Example. To illustrate our new approach and to compare it with collision integration schemes, we consider the simple example of a two-degree of freedom, mass-spring system, subject to a single ground constraint. See Figure 6.1 (b). We apply Störmer-Verlet as the base, unconstrained numerical method. Configuration is given by the positions of two mass particles, vertically oriented, such that the positions of the top and bottom mass particles are given by $\mathsf{q} = (q_1, q_2)^T \in \mathbb{R}^2$. The system is then subject to gravity and a ground-plane constraint on the bottom particle, $\mathsf{g}(\mathsf{q}) = q_2 \geq 0$. We set $m = 1$ for both mass particles. The linear spring stiffness is set to $k = 10$, spring length is $l = 5$, gravity is $g = 9.8$ and the base, unconstrained step-size is $h = 10^{-1}$.

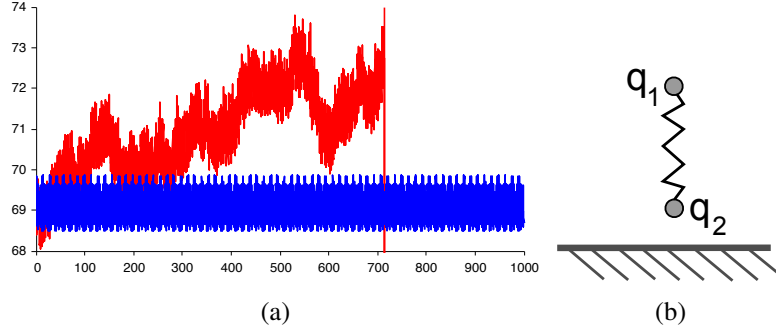In Figure 1.4(a) we plot in red the energy of the standard collision integration approach and in blue we

FIG. 6.1. *In (a) we plot the energy for the same simple, two-dimensional, two-point, mass-spring system, shown in (b), and discussed in Figure 1.4 and in §6.1.1 below. Here we use exactly the same physical parameters and algorithms as in the original example, with the exception that reflections for both the collision integrator and our extended reflection method are formulated to conserve Verlet's second-order numerical Hamiltonian, rather than the standard continuous Hamiltonian. We plot the energy for the first 1000 units of simulation time for both the collision integrator, in red, and our extended reflection scheme, in blue. Note that, after $t = 713$, the energy in the collision integration system blows up.*

plot the energy of our new extended reflection approach using a simple forward time-step predictor. In this first example both methods apply reflections that conserve the continuous Hamiltonian. The collision integration method, however, applies reflections at times of contact with the floor, while, in this simple example, our extended reflection method reduces to simply applying reflections at the beginning of all time-steps that would otherwise result in the bottom mass-particle passing through the ground-plane.

In the first time-period, shown in Figure 1.4(a), energy remains fully bounded for the extended reflection method proposed here and, in particular, varies in the expected manner of standard symplectic methods. Over the same period the hybrid solution experiences energy growth. After $t = 546$, energy in the collision integrator enters a highly oscillatory regime causing the solution to blow up. In Figure 1.4(b) we continue the energy plot of our extended reflection approach out to 100,000 units of simulation time.

Of course we are not limited to continuous Hamiltonian conserving reflections. We may, for instance, expect an improvement in behavior by applying reflections that more closely conserve the shadowing numerical Hamiltonian. In Figure 6.1 (a) we plot the energy of exactly the same two simulations discussed above *except*, in this example, we apply reflections that preserve Störmer-Verlet's second-order numerical Hamiltonian [Bond and Leimkuhler 2007]. We observe a small improvement in the bounds of the energy envelope for our extended reflection method. Similarly, we note that, although the collision integration scheme still experiences energy growth, the obtained solution does not blow up until later, at $t = 713$. Note that we could also consider applying higher order symplectic Gauss methods for unconstrained steps leading to collisions [Bond and Leimkuhler 2007]. Given the persistence of constraint events, however, this is effectively the same as simply switching to a higher-order, base, unconstrained integrator for both methods.

6.2. Discrete Limit Momenta. We next observe that applying the constraint force $\mathsf{N}(\mathsf{q}^t)\lambda$ in (5.5) with respect to the *next* time increment, $t + 1$, is indistinguishable from directly modifying the discrete momentum, p^t, at time t. We thus decompose the constraint multipliers λ into right components, $\lambda^+ = \lambda^{t^+}$, given by a previous nonsmooth reflection, and left components, $\lambda^- = \lambda^{(t+1)^-}$, given by the current discrete-smooth step. Applying an energy conserving, generalized reflection operator at the *end* of time-step t, we then correspondingly obtain discrete analogues of the left and right limit momenta,

$$\mathsf{p}^{t^-} = \mathsf{p}^t, \tag{6.1}$$

$$\mathsf{p}^{t^+} = \mathsf{p}^t + \mathsf{N}(\mathsf{q}^t)\lambda^+. \tag{6.2}$$

Here $\mathsf{N}(\mathsf{q}^t)\lambda^+$ gives the impulse generated by the reflection of p^{t^-}. Consistent with our discussion in §6.1, the above partition of λ requires generalized reflections to be applied exclusively at the beginning (or equivalently end) of each time-step. Given our above analysis, we then include the extended reflection in our definition of left and right discrete momenta.

At the beginning of each time step, we augment the active set with the set of all constraints that would be activated or violated by a predicted configuration. More concretely, we predefine a consistent *predictor*

14

configuration, q^p, and then define an *extended active set*,

$$\mathbb{A}(\mathsf{q}^t, \mathsf{q}^p) \stackrel{\text{def}}{=} \left\{ i: \ g_i(\mathsf{q}^p) \leq 0, \ i \in \{0, .., m\} \right\} \cup \left\{ j: \ g_j(\mathsf{q}^t) = 0, \ j \in \{0, .., m\} \right\}. \tag{6.3}$$

The corresponding *extended* discrete left and right momenta are then given by

$$\mathsf{p}^{t^-} = \mathsf{p}^t, \tag{6.4}$$

$$\mathsf{p}^{t^+} = \mathsf{p}^t + \mathsf{G}_{\mathbb{A}}(\mathsf{q}^t)\lambda^+. \tag{6.5}$$

6.3. Discrete Generalized Reflections. To compute the discrete reflection term given in (6.5) above, we require a discrete generalized reflection operator, **R**, of the form

$$\mathsf{p}^+ = \mathsf{p}^- + \mathsf{G}_{\mathbb{K}}(\mathsf{q})\lambda = \mathbf{R}(\mathsf{q}, \mathsf{p}^-, \mathbb{K}, E), \tag{6.6}$$

that defines a *unique* reflection with respect to an active constraint set $\mathbb{K}$. We require that any such reflection operator satisfy the discrete analogues of the *jump conditions* from §3.5, given by the discrete kinematic feasibility condition,

$$\mathsf{G}_{\mathbb{K}}(\mathsf{q})^T \mathsf{M}^{-1} \left(\mathsf{p}^- + \mathsf{G}_{\mathbb{K}}(\mathsf{q})\lambda \right) \geq 0, \tag{6.7}$$

the discrete Euler-Lagrange inclusion condition,

$$\mathsf{p}^+ - \mathsf{p}^- = \mathsf{G}_{\mathbb{K}}(\mathsf{q})\lambda, \quad \lambda \geq 0. \tag{6.8}$$

and conservation of a specified energy function, E, (e.g., the discrete, numerical, or continuous Hamiltonian) such that

$$E(\mathsf{q}, \mathsf{p}^+) = E(\mathsf{q}, \mathsf{p}^-). \tag{6.9}$$

As in the time-continuous case, whenever multiple active constraints are considered, i.e., $|\mathbb{K}| > 1$, this problem is underdetermined. From the infinitely many possible solutions, we select the one obtained by applying a discrete adaptation of the generalized reflection operator of Kaufman et al. [2010]. This operator simultaneously guarantees uniqueness, determinism, symmetry-preservation, and satisfies all jump conditions when multiple constraints are active. In §A we briefly derive this discrete extension and provide pseudo-code for the implementation we employ.

7. Generalized Discrete Setting. Returning to the full setting where both discrete-smooth and discrete-nonsmooth modes are possible, we finally consider a fully general DELI-based system. Starting with the DELI system given by (2.6), (2.7) and (2.8), recall that in §5 we presumed, *a priori*, the precondition of the left discrete-smoothness criteria to obtain the Discrete-Smooth Integrator given by (5.6) through (5.9). In the general setting, however, at any time a subset of the active constraints may be discrete-smooth, while their complement may be discrete-nonsmooth. In such cases we can expect that momentum will be negative with respect to some constraint gradients, and thus nonsmooth modes will be present.

To partition out constraint subsets that satisfy the discrete-smooth precondition at time t, we define the *smooth constraint set* as the set of discrete-smooth constraints, with respect to the discrete-right time-limit at t,

$$\mathbb{S}(\mathsf{q}^t, \mathsf{p}^{t^+}) \stackrel{\text{def}}{=} \left\{ i: \ \nabla g_i(\mathsf{q}^t)^T \mathsf{M}^{-1} \mathsf{p}^{t^+} = 0 \ \text{and} \ g_i(\mathsf{q}^t) = 0, \ i \in \{0, .., m\} \right\}. \tag{7.1}$$

Then, adding the discrete left and right jump terms given by (6.4) and (6.5), we include reflections and obtain a full, discrete, generalized extension of DELI for inequality-constrained systems,

$$\mathsf{p}^{t^+} = \mathsf{p}^t + \mathsf{G}_{\mathbb{A}}(\mathsf{q}^t)\lambda^+, \tag{7.2}$$

$$D_1 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}) + \mathsf{p}^{t^+} + \mathsf{N}_{\mathbb{S}}(\mathsf{q}^t)\lambda^- = 0, \tag{7.3}$$

$$0 \leq \lambda^- \perp \mathsf{g}_{\mathbb{S}}(\mathsf{q}^{t+1}) \geq 0, \tag{7.4}$$

$$\mathsf{p}^{t+1} = D_2 L_d(\mathsf{q}^t, \mathsf{q}^{t+1}) + \mathsf{N}_{\mathbb{S}}(\mathsf{q}^{t+1})\,\mu, \tag{7.5}$$

$$\mathsf{N}_{\mathbb{S}}(\mathsf{q}^{t+1})^T \mathsf{M}^{-1} \mathsf{p}^{t+1} = 0. \tag{7.6}$$

Algorithm 1 GVI(q, p, E, q^p)   // input: q^t, p^t, energy function, position predictor

1: $q^{old} \leftarrow q$	// cache current position
2: $\mathbb{A} \leftarrow \left\{ i : \ g_i(q^p) \leq 0, \ i \in \{0,..,m\} \right\} \cup \left\{ j : \ g_j(q) = 0, \ j \in \{0,..,m\} \right\}$	// compute extended active set
3: $p \leftarrow \mathbf{R}(q, p, \mathbb{A}, E)$	// apply generalized reflection
4: $\mathbb{S} \leftarrow \left\{ i : \ \nabla g_i(q)^T M^{-1} p = 0 \ \text{and} \ g_i(q) = 0, \ i \in \{0,..,m\} \right\}.$	// compute smooth set
5: $q \leftarrow \left\{ x : \ D_1 L_d(q, x) + p + N_{\mathbb{S}}(q)\lambda = 0, \ 0 \leq \lambda \perp g_{\mathbb{S}}(x) \geq 0 \right\}$	// apply position update
6: $p \leftarrow \left\{ y : \ y = D_2 L_d(q^{old}, q) + N_{\mathbb{S}}(q) \ \mu, \ N_{\mathbb{S}}(q)^T M^{-1} y = 0 \right\}$	// apply momentum update
7: **return** q, p	// output: q^{t+1}, p^{t+1}

7.1. Algorithm. With the above developments in place, our proposed discrete, generalized variational integrator (GVI) for fully generalized inequality constrained systems is then given by the pseudocode below in Algorithm 1. A GVI takes as input the current state, the selected energy function and the position predictor. It then computes the GVI step and outputs the next state.

7.2. Bilateral Constraints. Finally, we now further presume that, in addition to the set of inequality constraints, we also wish to enforce a set of equality constraints, as given by (1.1). When we consider equality constraints, we observe that the admissible set is now given by a projection of the inequality constraints onto the manifold defined by the additional equality constraints. In turn, this implies that normal cones are now given by the negative span of the active inequality constraint gradients, *projected down to the cotangent spaces* of the equality-constraint manifold.

It then remains to note that, in the generalized constraint setting, we treat equality constraints as active at all times. We then can add the equality constraints, f, and their corresponding gradients, which we will denote by F, to our discrete-smooth formulation, during all time-steps.

For our GVI methods this requires three basic changes. First, for arbitrary inequality constraint subsets, $\mathbb{K}$, the corresponding active inequality gradient subset must be projected onto the local manifold's cotangent space at q. Thus, when equality constraints are applied, we project[5] the active constraint gradient and redefine as

$$N_{\mathbb{K}}(q) \stackrel{\text{def}}{=} \left(I - \left(F(q)^T M^{-1} F(q)\right)^+ F(q) F(q)^T M^{-1} \right) G_{\mathbb{A}\mathbb{K}}(q), \quad \mathbb{A}\mathbb{K} = \{i : \ g_i(q) = 0, \ i \in \mathbb{K}\}. \quad (7.7)$$

Next, we simply modify the two constrained updates in the above pseudo-code, given in Algorithm 1, to include the equality constraints. We then obtain a new generalized equality-inequality position update step (to replace line 5 in the pseudocode):

$$q \leftarrow \left\{ x : \ D_1 L_d(q, x) + p + N_{\mathbb{S}}(q)\lambda + F(q)v = 0, \ 0 \leq \lambda \perp g_{\mathbb{S}}(x) \geq 0, \ f(x) = 0 \right\} \quad (7.8)$$

and a corresponding momentum update step (to replace line 6 in the pseudocode):

$$p \leftarrow \left\{ y : \ y = D_2 L_d(q^{old}, q) + N_{\mathbb{S}}(q)\mu + F(q)\xi, \ N_{\mathbb{S}}(q)^T M^{-1} y = 0, \ F(q)^T M^{-1} y = 0 \right\}. \quad (7.9)$$

Otherwise, the remaining pseudocode and all other previous discussion, derivation and guarantees (smooth-interval symplecticity, momentum conservation, etc.) continue to hold.

8. Further Numerical Examples. To understand and examine the behavior of the algorithms proposed here, we investigate implementations of GVI over a range of additional numerical examples. Although a wide variety of predictors can be employed, in the following examples we exclusively apply the simple, unconstrained, forward predictor,

$$q^p = \{x : \ D_1 L_d(q^t, x) + p^t = 0\}, \quad (8.1)$$

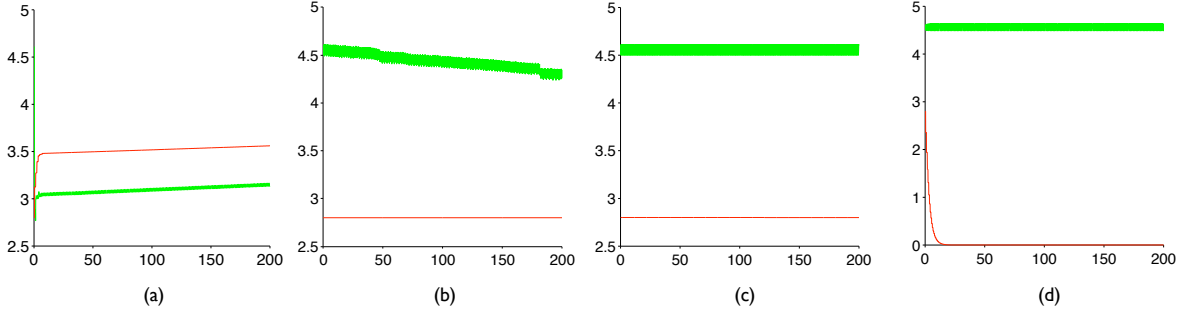[5]Here $\cdot^+$ indicates the standard matrix pseudoinverse.

FIG. 8.1. **Nonlinear Oscillator Example:** *In these figures we examine the behavior of various algorithms on the nonlinear oscillator example in two dimensions. All methods shown here begin with implicit midpoint as the base, unconstrained numerical method. See Figure 6.1 (c) and §8.0.1 for further details. Green lines and red lines plot energy and angular momentum respectively. Note that both energy and angular momentum should be preserved in this example. In* **(a)** *we plot the endpoint-constrained direct method. In* **(b)** *we plot the standard hybrid method. In* **(c)** *we plot the GVI integrator, where energy and angular momentum are both preserved long-term. Finally, in* **(d)**, *to highlight the importance of enforcing nonlinear constraints appropriately, we plot the GVI integrator using linearized constraints. Note that, for this last example, the GVI method continues to preserve energy; however, linearized constraints will no longer preserve rotational symmetry and thus, correspondingly, angular momentum is no longer preserved.*

and reserve an investigation of alternate predictors for later research. We implement implicit-midpoint and Störmer-Verlet based GVI-schemes using the implicit midpoint quadrature,

$$L_d(\mathsf{q}^k, \mathsf{q}^{k+1}) = hL\big(\frac{\mathsf{q}^{k+1} + \mathsf{q}^k}{2}, \frac{\mathsf{q}^{k+1} - \mathsf{q}^k}{h}\big), \tag{8.2}$$

and the Verlet quadrature,

$$L_d(\mathsf{q}^k, \mathsf{q}^{k+1}) = \frac{h}{2}\left(L(\mathsf{q}^k, \frac{\mathsf{q}^{k+1} - \mathsf{q}^k}{h}) + L(\mathsf{q}^{k+1}, \frac{\mathsf{q}^{k+1} - \mathsf{q}^k}{h}) \right), \tag{8.3}$$

respectively.

8.0.1. Nonlinear Oscillator. To examine a nonintregable system and momentum conservation properties, we next consider the example of a nonlinear oscillator in $\mathbb{R}^2$. We start with two spheres in the plane, with respective positions q_1 and q_2 in $\mathbb{R}^2$ and radii r_1 and r_2. The full configuration is then $\mathsf{q} = (\mathsf{q}_1, \mathsf{q}_2)^T$ with a corresponding nonlinear potential for the oscillator

$$V(\mathsf{q}) = \sum_{i=1}^{2} \| \mathsf{q}_i \|^2 (\| \mathsf{q}_i \|^2 - 1)^2. \tag{8.4}$$

See the inset figure below. We then impose the hard-sphere, non-overlap constraint between the two spheres using the nonlinear, inequality constraint

$$g(\mathsf{q}) = \| \mathsf{q}_1 - \mathsf{q}_2 \| - r_1 - r_2 \geq 0. \tag{8.5}$$

Note that both the potential *and* constraint in this example are invariant with respect to rotations and thus angular momentum should be conserved.

For this example all methods considered begin with implicit-midpoint as the base, unconstrained numerical method. We then set the step-size to $h = 10^{-1}$, mass and sphere radii are set to $m = 1$ and $r = 1$, respectively, for both particles. The starting configuration is given such that the particles are initially separated at the non-equilibrium position, $\mathsf{q}_1(0) = (0, -r - 4 \times 10^{-1})^T$, $\mathsf{q}_2(0) = (0, r + 4 \times 10^{-1})^T$, while the initial momentum initiates a counter-clockwise rotation with $\mathsf{p}_1(0) = (1, 0)^T$ and $\mathsf{p}_2(0) = (-1, 0)^T$. The system should then, as illustrated to the right, enter into a relative equilibrium composed of both periodic bouncing between the particles combined with a global rotation of the full system about the origin.

17
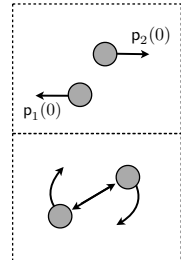
We compare the results of the endpoint-constrained direct method [6] (i.e., imposing $g(q^{t+1}) \geq 0$), the standard hybrid method, GVI applied using the full nonlinear constraint given in (8.5), and GVI applied using a linearization of the constraints. We include this last example to highlight the potential pitfalls associated with applying constraint linearization since this a popular simplifying strategy in the contact mechanics literature. In Figure 8.1 we plot the obtained energy in green and angular momentum in red for each method.

In Figure 8.1(a), as might be expected, the endpoint-constrained direct method quickly dissipates all normal modes of oscillation between the two spheres. After this initial dissipation, the spheres enter a relative equilibrium in which they orbit, in constant contact, with a purely smooth motion. This latter mode corresponds exactly to an endpoint-constrained direct method applied to a smooth system. As discussed in Section 3.1, direct methods applied to smooth motion are likewise not symplectic and thus are still not guaranteed to preserve energy. This is reflected in the plot, where the system continues to generate correspondingly poor energy behavior. In particular, for this latter phase, we see slow but continual energy growth.

In Figure 8.1(b) the standard collision integration method likewise displays poor energy behavior. As discussed in Section 1.3.2, energy drift (in the form of dissipation in this case) is generated, in collision integrators, by the interleaving of reflections with variable step-size, unconstrained integration steps. The reflection impulses, however, are applied along constraint gradients evaluated at the time of constraint activation and thus, despite the poor energetic behavior, momentum *is* conserved due to the rotational invariance of the constraint function.

In Figures 8.1(c) and (d) we plot GVI using the true nonlinear constraints and linearized variants respectively. Our GVI method, plotted in Figure 8.1(c), enforces the true nonlinear constraints and correspondingly we note that both approximate energy *and* exact angular momentum are conserved long-term, while the combined nonsmooth, rotational *and* oscillatory modes are all preserved. Finally, we note that, when applied to linearized constraints, as in Figures 8.1(d), GVI continues to preserve energy but, due to constraint linearization, rotational invariance no longer holds and thus angular momentum is no longer conserved. This highlights the cost of linearizing constraints when applying geometric methods.
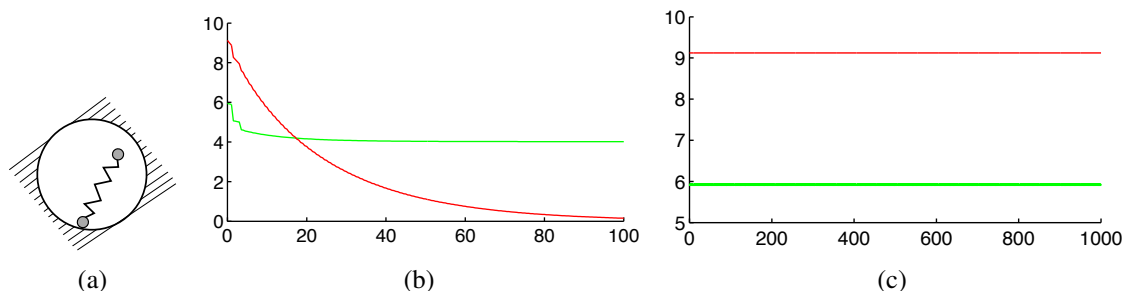

| (a) | (b) | (c) |

FIG. 8.2. *We consider the combined smooth and nonsmooth trajectory of a constrained system composed of the same simple constrained mass-spring system considered in §5.2 and Figure 5.1. Unlike the previous example, however, we now initialize the first particle to an off sphere boundary position and set the second particle to an on-boundary position as illustrated in (a). As in §5.2 both particles are initialized with corresponding unit length momenta that point counter-clockwise and tangent to the boundary. This initiates a counter-clockwise rotation for both particle with an oscillatory motion between them via forces applied by the spring. The first particle, however, should now begin a nonsmooth oscillatory bouncing off the surface, while the trajectory of the second particle should remain smooth and on boundary throughout the simulation. Otherwise all other physical parameters remain unchanged from the last example. In (b) we plot the energy (in green) and the angular momentum (in red) of the direct-substitution, end-point constrained, method for the first 100 units of simulation time. We note the dissipation in both angular momentum and energy. This includes a rapid decay of all nonsmooth bouncing modes. In (c) we similarly plot the (approximate) preservation of energy (in green) and (exact) conservation of angular momentum (in red) obtained by GVI out to 1000 units of simulation time. See §8.1 for further discussion.*

8.1. Smooth and Nonsmooth, Spring and Sphere. As a first consideration of an example with combined smooth *and* nonsmooth boundary modes, we reconsider our spring and sphere example from §5.2. In this example we leave all physical and time-step parameters unchanged from from §5.2 *except* that we now initialize the example system to a combined smooth *and* nonsmooth, rotational trajectory. We initialize the first particle to an off sphere boundary position at $q_1(0) = (4, -1)$. We then set the second particle to an on boundary position, $q_2(0) = (3, -4)$. See Figure 8.2(a). As in §5.2 both particles are initialized with corresponding unit length momenta that point counter-clockwise and tangent to the boundary. This initiates a counter-clockwise

---

[6]The midpoint-constrained direct method (i.e., imposing $g(\frac{q^{t+1}+q^t}{2}) \geq 0$) was highly unstable for this example.

rotation for both particles with an oscillatory motion between them via forces applied by the spring. The first particle, however, should now begin a nonsmooth oscillatory bouncing off the surface, while the trajectory of the second particle should remain smooth and on boundary throughout the simulation.

In Figure 8.2 (b) we plot the energy (in green) and the angular momentum (in red) of the direct, end-point constrained method for the first 100 units of simulation time. We note that decay in both angular momentum and energy for the direct method corresponds to the destruction of both the nonsmooth *and* smooth modes of the particles, as well as the rotational and oscillatory modes of the system. In Figure 8.2 (c) we similarly plot the (approximate) preservation of energy, in green, and the (exact) conservation of angular momentum, in red, obtained by our GVI method out to 1000 units of simulation time. Throughout this simulation, GVI maintains the nonsmooth bouncing mode of the first particle, the smooth-on-boundary mode of the second particle, as well as the rotational and oscillatory modes of the full system.
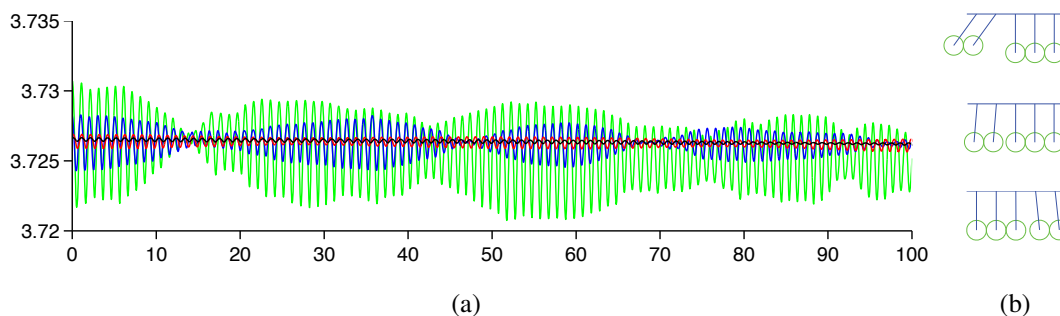

(a)   (b)

FIG. 8.3. **Newton's Cradle Example:** *In (a) we compare the results of our implicit-midpoint-based GVI method applied to a Newton's Cradle System over a range of decreasing step sizes, $h = 3 \times 10^{-2}$ (green), 2×10^{-2} (blue), 10^{-2} (red), and 5×10^{-3} (black). We note that in the passage from $h = 3 \times 10^{-2}$ to $h = 10^{-2}$, we obtain the classic transition to stable energy behavior that is characteristic of symplectic methods. In (b) we illustrate the behavior of the integrator with a series of snapshots from the simulation.*

8.1.1. Newton's Cradle. Next we examine an example with combined smooth *and* nonsmooth boundary modes *and* the simultaneous enforcement of both equality and inequality constraints. We consider the classic, pendulum-based, Newton's Cradle system composed of n initially touching spheres subject to gravity. Each sphere $i \in [1, n]$ is constrained by a corresponding equality constraint

$$f_i(q) = \| q_i - p_i \| - l = 0, \tag{8.6}$$

to remain at constant length l from its hanging attachment point p_i. Neighboring pairs i, j of spheres are then further constrained by non-overlap constraints

$$g_{i,j}(q) = \| q_i - q_j \| - r_i - r_j \geq 0. \tag{8.7}$$

Thus, for an n-sphere cradle we enforce n bilateral and $n - 1$ unilateral constraints and obtain a combination of smooth motion along the constraint boundary interfaces between neighboring spheres in moving and resting clusters, holonomic motion along the manifolds given by the pendulum constraints, and nonsmooth motion at each sphere impact.

For this example we begin with implicit-midpoint as the base, unconstrained numerical method. We then set the masses, sphere radii, and pendulum lengths to $m = 1$, $r = 0.25$ and $l = 1$, respectively, for all spheres. We examine a five-ball cradle. The starting configuration is given such that all but the two leftmost spheres are at rest and in contact. The two leftmost spheres are initially pulled back into a non-equilibrium, contacting position, at an angle of $-\pi/5$ from rest. The system should then enter into the characteristic Newton's cradle behavior as illustrated by the simulation snapshots shown in Figure 8.3 (b).

We compare the results of GVI over a range of decreasing step sizes. In Figure 8.3 (a) we plot the obtained energy for $h = 3 \times 10^{-2}$, 2×10^{-2}, 10^{-2}, and 5×10^{-3}. We note that in the passage from $h = 3 \times 10^{-2}$ to $h = 10^{-2}$ we obtain the classic transition to stable energy behavior that is characteristic of symplectic methods. Furthermore, as step size decreases, the cradle preserves the correct cyclical behavior, expected from Newton's cradle, for increasingly longer periods. Finally, we note that both the direct methods, and the hybrid schemes, along with generating their expected energy errors, do not predict the correct physical behavior of Newton's Cradle.
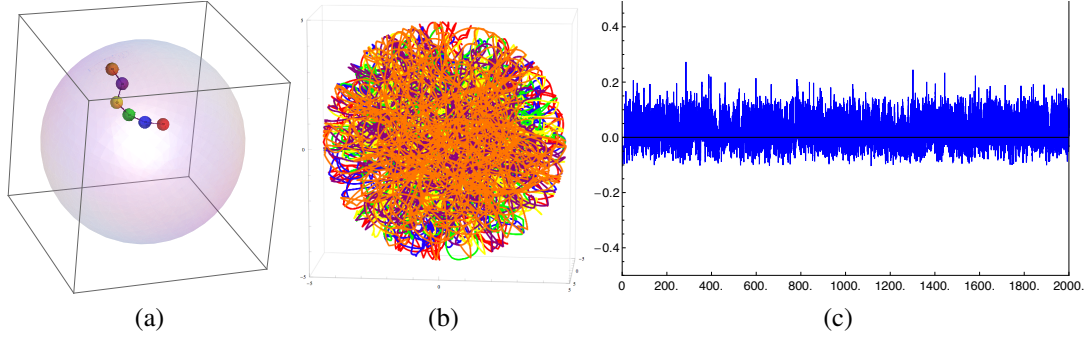
FIG. 8.4. **Lennard-Jones Example:** *In these figures we examine the behavior of Störmer-Verlet-based GVI on a system of six particles in $\mathbb{R}^3$ connected sequentially by rods and interacting in a Lennard-Jones potential. In* **(a)** *we take a snapshot of the system configuration,* **(b)** *traces out the entire trajectory of all particle centers over* 2000 *units of simulation time, stepped at $h = 10^{-2}$, while* **(c)** *plots the change in energy over the same trajectory.*

8.1.2. Interacting Constrained Particles under Lennard-Jones. Finally, motivated by potential applications in Molecular Dynamics [Frenkel and Smit 2001], we consider the example of six spheres in $\mathbb{R}^3$, i.e., $q = (q_1^T, ..., q_6^T)^T \in \mathbb{R}^{18}$, connected sequentially by bilateral rod constraints,

$$f_i(q) = \| q_i - q_{i+1} \| - l = 0. \tag{8.8}$$

Sphere-pairs are constrained by non-overlap conditions,

$$g_{i,j}(q) = \| q_i - q_j \| - 2r \geq 0, \tag{8.9}$$

interact in a Lennard-Jones potential,

$$V(q) = \sum_{i \neq j} 4\varepsilon \left[\left(\frac{\sigma}{\| q_i - q_j \|} \right)^{12} - \left(\frac{\sigma}{\| q_i - q_j \|} \right)^6 \right], \tag{8.10}$$

and are further constrained to lie in the interior of a bounding sphere of radius r_B so that additionally,

$$g_{i,b}(q) = - \| q_i \| - r + r_b \geq 0. \tag{8.11}$$

In our simulation we employ reduced time units, $(m\sigma^2/\varepsilon)^{1/2}$, with the mass of all particles set to unity, $\sigma/2 = r = 0.3$, $r_b = 5$, and $l = 1$. Figure 8.1.1 (a) shows a snapshot of the system configuration. Figure 8.1.1 (b) traces out the entire trajectory computed by Störmer-Verlet-based GVI of all particle centers over 2000 units of simulation time, stepped at $h = 10^{-2}$, while Figure 8.1.1 (c) plots the change in energy over the same trajectory.

9. Discussion and Conclusion. We have introduced a fully nonsmooth, discrete Hamilton's Principle and an associated DELI integration formulation that enables the generation of VI-based, geometric methods for the numerical integration of generalized inequality/equality-constrained, nonsmooth Hamiltonian systems. Adding additional structure to this framework we have obtained the first family of DELI-based integration methods: GVIs. By construction the GVI methods proposed here preserve momentum and equality constraints. They enforce simultaneous inequality constraints, obtain smooth unilateral motion along constraint boundaries and allow for both nonsmooth and smooth boundary approach and exit trajectories. Finally, we have shown that the proposed GVIs are symplectic over smooth-trajectory intervals and are observed to conserve energy approximately throughout.

We have validated these properties in our numerical experiments and compared them against corresponding direct-substitution and collision-integration schemes in the literature. We have shown that our proposed methods resolve systems subject to multiple persistent, frequent and/or simultaneously active constraints. We have further tested these methods on difficult scenarios, where both smooth and nonsmooth active constraint modes occur simultaneously with high frequency; generally, in many of the examples presented above, during the majority of time steps.

In many of the above examples, potentially difficult, nonlinear, numerical optimization problems were solved. In particular, we note that the complementarity conditions in our Discrete Smooth Integrator are,

20

unlike more standard KKT-type complementarity conditions, asymmetric with respect to constraint and constraint gradient evaluations. As such, equations (5.6) – (5.9), do not correspond to the optimality conditions of standard nonlinear optimization problems [Bertsekas 1999] and instead require customized numerical solution strategies. In this work we have focused on investigating the challenges of inequality-constrained, numerical integration and then on developing and validating our proposed methods. In future work we will explore efficient computational techniques for the scalable solution of these DELI-derived optimization problems. In practice, however, for the numerical experiments discussed here, we have found that relatively simple customized approaches were possible that allowed for efficient solutions, while preliminary experiments investigating the possibility of extending these approaches to larger-scale problems are promising.

Further, we look forward to examining the potential trade-offs between energy behavior and various degrees of constraint enforcement guarantees. We also expect that a rigorous constraint error analysis, based on choice of predictor, should be a useful complement to the GVI methods proposed here. Finally, we note that GVI methods are the first of many possible instantiations of DELI methods. The geometric integration of inequality-constrained systems clearly remains a challenging area of investigation and we hope that the DELI framework will lead to further fruitful explorations, and the development of new complementary numerical methods.

References.

Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999.

Stephen D. Bond and Benedict J. Leimkuhler. Stabilized integration of Hamiltonian systems with hard-sphere inequality constraints. *SIAM Journal on Scientific Computing*, 30:134–147, 2007.

Bernard Brogliato. *Nonsmooth Mechanics*. Springer-Verlag, 1999.

A. Chatterjee and A. L. Ruina. A new algebraic rigid-body collision law based on impulse space considerations. *Journal of Applied Mechanics*, 65(4):939–951, 1998.

Fehmi Cirak and Matthew West. Decomposition contact response (DCR) for explicit finite element dynamics. *International Journal for Numerical Methods in Engineering*, 64(8):1078–1110, 2005.

F. H. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley and Sons, 1983.

Peter Deuflhard, Rolf Krause, and Susanne Ertel. A contact-stabilized newmark method for dynamical contact problems. *International Journal for Numerical Methods in Engineering*, 73:1274–1290, 2008.

David Doyen, Alexandre Ern, and Serge Piperno. Time-integration schemes for the finite-element dynamic Signorini problem. *SIAM J. Sci. Comp.*, 33:223–249, 2011.

R. C. Fetecau, J. E. Marsden, M. Ortiz, and M. West. Nonsmooth Lagrangian mechanics. *SIAM Journal on Applied Dynamical Systems*, 2:381–416, 2003.

D. Frenkel and B. Smit. *Understanding Molecular Simulation,*. Academic Press, 2001.

Christof Glocker. Concepts for modeling impacts without friction. *Acta Mechanica*, 168:1–19, 2004.

E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration: Structure-Preserving Algorithms for Odinary Differential Equations*. Springer, 2002.

D. Heyes. Molecular dynamics simulations of restricted primitive model 1:1 electrolytes. *Chemical Physics*, 69:153–163, 1982.

Yao A. Houndonougbo, Brian B. Laird, and Benedict J. Leimkuhler. A molecular-dynamics algorithm for mixed hard-core/continuous potentials. *Molecular Physics*, 98:309–316, 2000.

C. Kane, E. Repetto, M. Ortiz, and J. Marsden. Finite element analysis of nonsmooth contact. *Comp. Meth. Appl. Mech. Engrg. 180 (1999)*, 1999.

Danny M. Kaufman, Eitan Grinspun, and Dinesh K. Pai. Generalized reflections and multiimpact systems. 2010.

Houari Boumediène Khenous, Patrick Laborde, and Yves Renard. Mass redistribution method for finite element contact problems in elastodynamics. *European Journal of Mechanics A/Solids*, 27:918–932, 2008.

N. Kikuchi and J.T. Oden. *Contact Problems in Elasticity: A study of variational inequalities and finite element methods*. SIAM Studies in Applied and Numerical Mathematics, 1988.

Rolf Krause and Mirjam Walloth. Presentation and comparison of selected algorithms for dynamic contact based on the newmark scheme. *ICS Preprint NO. 2009-08*, 2009.

T. A. Laursen and V. Chawla. Design of energy conserving algorithms for frictionless dynamic contact problems. *IJNME*, 40:863–886, 1997.

T. A. Laursen and G. R. Love. Improved implicit integrators for transient impact problems—geometric admissibility within the conserving framework. *IJNME*, 53:245–274, 2002.

Benedict Leimkuhler and Sebastian Reich. *Simulating Hamiltonian Dynamics*. Cambridge, 2004.

J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, pages 357–514, 2001.

William J. McNeil and William G. Madden. A new method for the molecular dynamics simulation of hard core molecules. *The Journal of Chemical Physics*, 76(12):6221–6226, 1982.

J. J. Moreau. Unilateral contact and dry friction in finite freedom dynamics. *Nonsmooth Mechanics and Applications, CISM Courses and Lectures*, (302):1–82, 1988.

A. Pandolfi, C. Kane, J. E. Marsden, and M. Ortiz. Time-discretized variational formulation of non-smooth frictional contact. *International Journal for Numerical Methods in Engineering*, 53(8):1801–1829, 2002.

R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

Ralph Tyrrell Rockafellar and Roger J.B. Wets. *Variational Analysis*. Springer, 1998.

J.M. Sanz-Serna and M.P. Calvo. *Numerical Hamiltonian Problems*. Chapman and Hall, 1994.

D. E. Stewart. Rigid-body dynamics with friction and impact. *SIAM Rev.*, 42(1):3–39, 2000.

Richard M. Stratt, Stephen L. Holmgren, and David Chandler. Constrained impulsive molecular dynamics. *Molecular Physics*, 42:1233–1143, 1981.

S. H. Suh, L. Mier y Teran, H. S. Whitea, and H. T. Davis. Molecular dynamics study of the primitive model of 1–3 electrolyte solutions. *Chemical Physics*, 142:201–211, 1990.

M. West. *Variational Integrators*. PhD thesis, California Institute of Technology, 2004.

Appendix A. The Discrete Generalized Reflection Operator.

For an arbitrary choice of energy function, $E(q, p)$, we invoke the variational model from Kaufman et al. [2010], in which a generalized reflection operator is interpreted as a sequence of energy preserving, constraint satisfying projections. Following their time-continuous case derivation we then generate the corresponding discrete generalized reflection analogue as the discrete, energy preserving, constraint satisfying projection sequence defined in Algorithm 2 below.

Algorithm 2 `Discrete_Generalized_Reflection(q,p,`$\mathbb{K}$`,`E`)`

1: **while** true **do**
2: $\quad$ $\mathbb{V} \leftarrow \emptyset$
3: $\quad$ **for** k in $\mathbb{K}$ **do**
4: $\quad\quad$ **if** $\nabla g_k(q)^T M^{-1} p < 0$ **then**
5: $\quad\quad\quad$ $\mathbb{V} \leftarrow k$
6: $\quad\quad$ **end if**
7: $\quad$ **end for**
8: $\quad$ **if** $\mathbb{V} \neq \emptyset$ **then**
9: $\quad\quad$ $\lambda \leftarrow \left\{ y : \ E(q, p + G_{\mathbb{V}}(q)y) = E(q,p), \ G_{\mathbb{V}}(q)^T M^{-1}\big(p + G_{\mathbb{V}}(q)y\big) \geq 0, \ y \geq 0 \right\}$
10: $\quad\quad$ $p \leftarrow p + G_{\mathbb{V}}(q)\lambda$
11: $\quad$ **else**
12: $\quad\quad$ **return** p
13: $\quad$ **end if**
14: **end while**

THEOREM A.1. *At termination, Algorithm 2 satisfies all discrete jump conditions.*

Proof. By construction, each iteration satisfies conditions (6.8) and (6.9). Termination guarantees condition (6.7) . □

When the energy function is given by the separable continuous Hamiltonian, e.g., $E = H$, the discrete generalized reflection operator is identical to the time-continuous case considered in Kaufman et al. [2010]. In this case each solve of line 9, in the above pseudo-code, is obtained by the non-negative, least-squares solution of the normal equation:

$$G_{\mathbb{V}}(q)^T M^{-1} G_{\mathbb{V}}(q)\lambda = -2G_{\mathbb{V}}(q)^T M^{-1} p, \quad \lambda \geq 0. \tag{A.1}$$

More generally, whenever the Energy function is quadratic in p, e.g., as in the Störmer-Verlet's second-order numerical Hamiltonian [Bond and Leimkuhler 2007], the solution to line 9 is likewise given by a non-negative, linear least-squares system. Finally, for possible energy functions containing higher order polynomials in p, more complex, variational treatments are required.

Finally, we note that, as discussed above, GVI is agnostic to the choice of reflection model *as long as* a *well-posed* solution exists that satisfies the discrete jump conditions, (6.7) - (6.9). As in the case of the Generalized Reflection operator, alternate existing multi-impact models can be similarly discretized to do so. As a particularly compact example, Moreau's elastic multi-impact solution [Moreau 1988] can be discretized to obtain the energy preserving discrete reflection given in Algorithm 3 below.

Algorithm 3 Discrete_Moreau_Reflection$(\mathsf{q},\mathsf{p},\mathbb{K},E)$

1: $\lambda \leftarrow \left\{ \mathsf{y} : \ E(\mathsf{q},\mathsf{p}+\mathsf{G}_\mathbb{K}(\mathsf{q})\mathsf{y}) = E(\mathsf{q},\mathsf{p}), \ \mathsf{G}_\mathbb{K}(\mathsf{q})^T\mathsf{M}^{-1}\big(\mathsf{p}+\mathsf{G}_\mathbb{K}(\mathsf{q})\mathsf{y}\big) \geq 0, \ \mathsf{y} \geq 0 \right\}$

2: $\mathsf{p} \leftarrow \mathsf{p}+\mathsf{G}_\mathbb{K}(\mathsf{q})\lambda$

3: **return** p

Appendix B. Direct Substitution Schemes.

As discussed above, an extended value potential for inequality constrained systems can be composed by concatenating the unconstrained Hamiltonian system's potential with the extended value indicator function. Following this observation, the *direct-substitution* approach in symplectic methods [Kane et al. 1999; Stewart 2000; Pandolfi et al. 2002; Deuflhard et al. 2008] is to compose direct extensions by treating the nonsmooth constraint force, $-\partial I_\mathsf{A}(\mathsf{q})$, as simply an additional force to be handled by each numerical method in the usual manner.

B.1. Direct Integration Examples. To illustrate this approach, in the following, we will present examples using two popular integration approaches: the implicit midpoint method and the Newmark family of integration algorithms.

B.1.1. Implicit Midpoint Method. The implicit midpoint method is given by the discrete, momentum, phase-space pair

$$\mathsf{q}^{t+1} = \mathsf{q}^t + \frac{h}{2}\mathsf{M}^{-1}(\mathsf{p}^{t+1} + \mathsf{p}^t), \tag{B.1}$$

$$\mathsf{p}^{t+1} = \mathsf{p}^t - h\nabla V(\frac{\mathsf{q}^{t+1} + \mathsf{q}^t}{2}). \tag{B.2}$$

A direct-substitution, nonsmooth-constrained extension to the midpoint rule is generated by adding the extended value indicator to the unconstrained system's potential. Applying a generalized gradient, this gives

$$\mathsf{q}^{t+1} = \mathsf{q}^t + \frac{h}{2}\mathsf{M}^{-1}(\mathsf{p}^{t+1} + \mathsf{p}^t), \tag{B.3}$$

$$\mathsf{p}^{t+1} \in \mathsf{p}^t - h\nabla V(\frac{\mathsf{q}^{t+1} + \mathsf{q}^t}{2}) - h\partial I_\mathsf{A}(\frac{\mathsf{q}^{t+1} + \mathsf{q}^t}{2}). \tag{B.4}$$

An alternate, end-point constrained variant of implicit midpoint can be obtained by simply evaluating constraint forces at the end of the time-step [Stewart 2000]. This leads to the modified midpoint method momentum update

$$\mathsf{p}^{t+1} \in \mathsf{p}^t - h\nabla V(\frac{\mathsf{q}^{t+1} + \mathsf{q}^t}{2}) - h\partial I_\mathsf{A}(\mathsf{q}^{t+1}). \tag{B.5}$$

B.1.2. Newmark Methods. The Newmark family of integration schemes are given by the discrete, velocity, phase-space pair

$$\mathsf{q}^{t+1} = \mathsf{q}^t + h\dot{\mathsf{q}}^t - \frac{h^2}{2}\Big[(1-2\beta)\,\mathsf{M}^{-1}\nabla V(\mathsf{q}^t) + 2\beta\,\mathsf{M}^{-1}\nabla V(\mathsf{q}^{t+1})\Big], \tag{B.6}$$

$$\dot{\mathsf{q}}^{t+1} = \dot{\mathsf{q}}^t - h\Big[(1-\gamma)\,\mathsf{M}^{-1}\nabla V(\mathsf{q}^t) + \gamma\,\mathsf{M}^{-1}\nabla V(\mathsf{q}^{t+1})\Big]. \tag{B.7}$$

Recall that Newmark methods vary $\beta \in [0, \frac{1}{2}]$ and $\gamma \in [0,1]$. Newmark is second order accurate if and only if $\gamma = \frac{1}{2}$, otherwise it is only consistent. Implicit Newmark with $\beta = \frac{1}{4}$ and $\gamma = \frac{1}{2}$ gives the implicit trapazoidal method, while $\beta = 0$ and $\gamma = \frac{1}{2}$ gives explicit Newmark.

A direct-substitution, nonsmooth-constrained extension of the Newmark Family is generated by adding the extended value indicator to the unconstrained system's potential. Applying a generalized gradient, this gives

$$q^{t+1} \in q^t + h\dot{q}^t - \frac{h^2}{2}\left[(1-2\beta)\,M^{-1}\left(\nabla V(q^t)+\partial I_A(q^t)\right)+2\beta\,M^{-1}\left(\nabla V(q^{t+1})+\partial I_A(q^{t+1})\right)\right], \quad (B.8)$$

$$\dot{q}^{t+1} \in \dot{q}^t - h\left[(1-\gamma)\,M^{-1}\left(\nabla V(q^t)+\partial I_A(q^t)\right)+\gamma\,M^{-1}\left(\nabla V(q^{t+1})+\partial I_A(q^{t+1})\right)\right]. \quad (B.9)$$

An Implicit/Explicit (IMEX) variant of this extension was developed by Kane et al. [1999] in which both β and γ are enforced as fully implicit values (i.e., $\beta = \frac{1}{2}, \gamma = 1$) for *just* the nonsmooth portion of the composite potential. This generates a family of nonsmooth IMEX Newmark integrators given by

$$q^{t+1} \in q^t + h\dot{q}^t - \frac{h^2}{2}\left[(1-2\beta)\,M^{-1}\nabla V(q^t)+2\beta\,M^{-1}\nabla V(q^{t+1})\right]-\frac{h^2}{2}M^{-1}\partial I_A(q^{t+1}), \quad (B.10)$$

$$\dot{q}^{t+1} \in \dot{q}^t - h\left[(1-\gamma)\,M^{-1}\nabla V(q^t)+\gamma\,M^{-1}\nabla V(q^{t+1})\right]-hM^{-1}\partial I_A(q^{t+1}). \quad (B.11)$$

B.2. Variational Interpretation of Direct Nonsmooth Integrators: Optimization Forms. Given the explicitly variational definition of the generalized gradient [Rockafellar and Wets 1998], direct, nonsmooth constrained extensions of symplectic methods, as described above, generally reduce to constrained, nonlinear minimizations of the form

$$q^{t+1} = \operatorname*{argmin}_{q}\left\{ e(q^t,q,h) : f(q^t,q,h) \in A \right\}. \quad (B.12)$$

Interpreted in this equivalent form, symplectic methods, applied to Hamiltonian systems, generate an objective function, e, generally polynomial in q, and a vector-valued, constrained configuration function, f, linear in q.

B.2.1. Implicit Midpoint Example. As an example, consider the nonsmooth midpoint rule. We substitute Equation (B.4) into (B.3) to obtain the DI

$$q^{t+1} \in q^t + hM^{-1}p^t - \frac{h^2}{2}M^{-1}\nabla V(\frac{q^t+q^{t+1}}{2}) - \frac{h^2}{2}M^{-1}\partial I_A(\frac{q^t+q^{t+1}}{2}). \quad (B.13)$$

Applying the definition of the generalized gradient, we then obtain an equivalent local, nonlinear minimization that corresponds to the template given by (B.12),

$$q^{t+1} = \operatorname*{argmin}_{q}\left\{\frac{1}{2}q^T M q - q^T\left(q^t + hM^{-1}p^t\right) + h^2 V(\frac{q^t+q}{2}) : (\frac{q^t+q}{2}) \in A\right\}. \quad (B.14)$$

Rearranging (B.3) the corresponding midpoint method, momentum update is then given by

$$p^{t+1} = \frac{2}{h}M(q^{t+1}-q^t) - p^t. \quad (B.15)$$

Newmark and other direct methods follow similarly.

B.3. Direct Method Behavior. Unfortunately, many of the characteristic advantages of symplectic methods are lost in the above direct nonsmooth treatment. The chief observation is, in the smooth setting, the good behavior of symplectic methods is ascribed, via backwards error analysis, to the existence of trajectory shadowing *numerical* Hamiltonians [Hairer et al. 2002]. More specifically, the flow of fixed-step, symplectic methods follow, up to truncation, a numerical Hamiltonian system that is constructed via a series expansion. Because the numerical Hamiltonian is obtained under a smoothness assumption, which does not hold in the nonsmooth case, standard backwards error analysis guarantees no longer hold for the inequality constrained systems [Stewart 2000; Bond and Leimkuhler 2007].

Despite these issues, during free motion (i.e., when constraints are inactive), these nonsmooth constrained symplectic methods trivially reduce to their corresponding unconstrained method and are thus still guaranteed to shadow numerical Hamiltonians. At each constraint boundary event, however, this guarantee is lost, and the solution obtained no longer matches the shadow trajectory.

In practice, it is exactly at these boundary events that such direct methods obtain a wide variety of incorrect and undesirable behaviors. Constraint enforcement generates dissipative trajectories for some direct methods, energy growth in others, and in many cases produces effectively nondeterministic restitution behavior both when varying time-step sizes across simulations and within a single fixed step-size simulation [Stewart 2000].

B.4. No Energy Conservation. Considering (B.12) we, in particular, note that, perhaps surprisingly, a large reason for these difficulties stems directly from the underlying variational structure of these methods.

THEOREM B.1. *All direct, one-step, nonsmooth symplectic methods that can be posed in the form of (B.12), can not, in general, preserve, either approximately, nor exactly, the energy of the underlying unconstrained symplectic method.*

Proof. For all such methods, (B.12) can be rewritten, with respect to the constraint functions, as

$$\mathsf{q}^{t+1} = \underset{\mathsf{q}}{\operatorname{argmin}} \left\{ e(\mathsf{q}^t, \mathsf{q}, h) : \mathsf{g}\Big((1-\alpha)\mathsf{q}^t + \alpha\mathsf{q}\Big) \geq 0 \right\}, \tag{B.16}$$

for some fixed $\alpha \in (0, 1]$.

We then note that we retrieve the solution of the underlying, unconstrained symplectic method by solving for q in

$$\nabla_{\mathsf{q}} e(\mathsf{q}^t, \mathsf{q}, h) = 0. \tag{B.17}$$

The constrained variant will return this solution for any step in which the unconstrained solution returns a non-violating configuration. More generally, formulating the Karush-Kuhn-Tucker (KKT) first order optimality conditions [Bertsekas 1999], for the above minimization, we obtain, for optimal $\mathsf{q}^{t+1} = \mathsf{q}$,

$$\nabla_{\mathsf{q}} e(\mathsf{q}^t, \mathsf{q}, h) - \alpha \nabla \mathsf{g}\Big((1-\alpha)\mathsf{q}^t + \alpha\mathsf{q}\Big)\lambda = 0, \tag{B.18}$$

$$0 \leq \lambda \perp \mathsf{g}\Big((1-\alpha)\mathsf{q}^t + \alpha\mathsf{q}\Big) \geq 0. \tag{B.19}$$

We then consider any time step in which an unconstrained solution violates constraints. The above KKT system implies that constraint forces are applied along constraint gradients and, more specifically, are of the form $\alpha \nabla \mathsf{g} \lambda, \ \lambda \geq 0$.

The complementarity term of the above KKT system then additionally requires, component-wise, that a constraint force magnitude, applied along the gradient ∇g_i, can be nonzero (and thus enforce the corresponding constraint, g_i) if and only if

$$g_i\Big((1-\alpha)\mathsf{q}^t + \alpha\mathsf{q}^{t+1}\Big) = 0. \tag{B.20}$$

Thus constraining forces can only be applied along constraint gradients (to enforce feasibility) if the position at which the constraint force is evaluated, $(1-\alpha)\mathsf{q}^t + \alpha\mathsf{q}^{t+1}$, lies *exactly* on the corresponding constraint's boundary. Since these methods demand constraint enforcement by construction, they implicitly enforce this additional condition as a secondary constraint. In turn, this implicit constraint overrides the underlying symplectic structure of the base, unconstrained method. In particular, the magnitude of the corresponding constraint force, that achieves this secondary constraint, scales independently of time-step size and energy. Instead, these constraint force magnitudes scale with constraint geometry and state, so that an arbitrarily large perturbation may be applied to the discrete system's energy and momentum in order to place the configuration at which constraint forces are evaluated, $(1-\alpha)\mathsf{q}^t + \alpha\mathsf{q}^{t+1}$, on the admissible set boundary. □

Appendix C. The Discrete-Smooth Integrator. We now show symplecticity and momentum conservation for the Discrete-Smooth Integrator (DSI) by proving Theorem 5.1.

Proof. As a preliminary, starting at time t^+, we first define an *Oracle Set* composed of constraint indices that will be active at the end of the current time step, $\mathbb{O}(\mathsf{q}^{t+1}) \overset{\text{def}}{=} \{i : g_i(\mathsf{q}^{t+1}) = 0\}$. The intersection of the oracle set with the *Smooth Set*, $\mathbb{S}(\mathsf{q}^t, \mathsf{p}^{t^+})$, defined in (6.1), gives $\mathbb{I} \overset{\text{def}}{=} \mathbb{O}(\mathsf{q}^{t+1}) \cap \mathbb{S}(\mathsf{q}^t, \mathsf{p}^{t^+})$.

Now consider the optimal λ^- and μ that obtain (q^{t+1}, p^{t+1}) in (6.3) – (6.6). We then observe that setting $\gamma = \lambda^-$ and $\delta = \mu$ we also satisfy the following equality constrained integration system

$$D_1 L_d(q^t, q^{t+1}) + p^t + G_{\mathbb{I}}(q^t)\gamma = 0, \tag{C.1}$$

$$g_{\mathbb{I}}(q^{t+1}) = 0, \tag{C.2}$$

$$p^{t+1} = D_2 L_d(q^t, q^{t+1}) + G_{\mathbb{I}}(q^{t+1})\,\delta, \tag{C.3}$$

$$G_{\mathbb{I}}(q^{t+1})^T M^{-1} p^{t+1} = 0. \tag{C.4}$$

In particular, plugging $\gamma = \lambda^-$ and $\delta = \mu$ into the above system gives *exactly* the same (q^{t+1}, p^{t+1}) solution as the DSI system.

Next we observe that Equations (5.6) through (5.9) give the position-momentum form of the equality constrained discrete Hamiltonian map of Marsden and West [2001, (3.5.2a) – (3.5.2d)]. Symplecticity and momentum conservation then follow directly by noting that restricting DSI steps to constraints indexed in $\mathbb{S}(q^t, p^{t^+})$ guarantees initial position and momentum feasibility [Marsden and West 2001, (3.5.1)] for the corresponding equality constrained discrete Hamiltonian maps, at all steps. $\square$